THE BULLETIN OF
Mathematical
BIOPHYSICS

THE UNIVERSITY OF CHICAGO PRESS · CHICAGO · ILLINOIS

VOLUME 14

1952

PUBLISHED MARCH, JUNE, SEPTEMBER, AND DECEMBER, 1952
PRINTED BY THE UNIVERSITY OF CHICAGO PRESS, CHICAGO, ILLINOIS

THE BULLETIN OF Mathematical BIOPHYSICS

MARCH 1952

THE UNIVERSITY OF CHICAGO PRESS · CHICAGO

VOLUME 14 · · · · · NUMBER 1

The Bulletin of
MATHEMATICAL BIOPHYSICS

Editor:

N. RASHEVSKY

Associate Editors:

H. D. LANDAHL *and* ANATOL RAPOPORT

The BULLETIN is devoted to publications of research in Mathematical Biology, as described on the inside back cover.

THE BULLETIN is published by the University of Chicago at the University of Chicago Press, 5750 Ellis Avenue, Chicago 37, Illinois, quarterly, in March, June, September, December. ¶The subscription price is $7.50 per year, the price of single copies is $2.25. Orders for service of less than a full year will be charged at the single-copy rate. ¶Patrons are requested to make all remittances payable to the University of Chicago Press in postal or express money orders or bank drafts.

THE FOLLOWING is an authorized agent:

For the British Empire, except North America, India, and Australasia: The Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1. Prices of yearly subscriptions and of single copies may be had on application.

CLAIMS FOR MISSING NUMBERS should be made within the month following the regular month of publication. The publishers expect to supply missing numbers free only when losses have been sustained in transit, and when the reserve stock will permit.

BUSINESS CORRESPONDENCE should be addressed to the University of Chicago Press, Chicago 37, Ill.

COMMUNICATIONS FOR THE EDITOR and manuscripts should be addressed to N. Rashevsky, Editorial Office of *The Bulletin of Mathematical Biophysics*, 5741 Drexel Avenue, Chicago 37, Ill.

NOTICE TO SUBSCRIBERS

If you change your address, please notify us and your local postmaster immediately. The Post Office does not forward third-class mail.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

A MATHEMATICAL ANALYSIS OF CARBON DIOXIDE RESPIRATION IN MAN*

ARTHUR B. CHILTON AND RALPH W. STACY
COMMANDER, CIVIL ENGINEER CORPS, U.S. NAVY AND
DEPARTMENT OF PHYSIOLOGY, OHIO STATE UNIVERSITY

By means of physico-mathematical models, formulas are obtained for the purpose of making analyses of external carbon dioxide respiration in man under conditions of metabolic equilibrium. The methods evolved are flexible, permitting study of a wide variety of physiological assumptions. Sample calculations relate various factors in the process of external respiration, and show good agreement with experimental data. A quantitative description of the effects of various factors on rate of carbon dioxide output and alveolar tension is given. In particular, this theory gives a quantitative prediction of the fluctuation in alveolar carbon dioxide tension over the period of a single respiratory cycle.

I. *Introduction.* Until recently physiologists have been unable to obtain instantaneous values of gaseous tensions within the alveolar spaces of the lungs and associated blood vessels because of sampling and instrumental difficulties. Recent improvements in techniques and instruments (Hunter, et al., 1948; Stacy, et al., 1948; Hitchcock and Stacy, 1950) now enable better understanding of dynamic changes occurring during this cycle. It seems appropriate that a broad theoretical attack be made on the subject to aid in understanding the processes involved.

This paper initiates such an attack by presenting a logical analysis, semi-empirical in nature, of the kinetics of carbon dioxide respiration in the lungs under conditions of metabolic equilibrium. To do so, physico-mathematical models are set up to represent the respiratory system. Actual conditions have been simplified to permit mathematical analysis. This simplification has not destroyed applicability of theoretical results to practical goals.

To avoid misunderstanding, we wish to emphasize that this paper does not purport to solve the problems of CO_2 respiration completely. The phrase "metabolic equilibrium," which we used above, refers to a condition of the human organism in which all the factors involved in respiration

* The work described in this paper was performed in part under contract between the Office of Naval Research and the Ohio State University Research Foundation.

have reached a level of adjustment with each other (and with the external environment) which permits the breathing cycle to be completely repetitive in nature. In our analysis, therefore, we may consider the physical factors of respiration rate, cardiac output, carbon dioxide level of venous blood, et cetera, to be established and constant. More extensive analysis, wherein these values are not in an equilibrium condition, requires a treatment of their inter-relationships and is left as a project for the future.

Under these conditions, it is proposed herein to use the models to (a) analyze carbon dioxide tensions assuming constant alveolar volume, (b) generalize to cases of inspiration and expiration, (c) determine cyclic changes in tension within the alveoli, and (d) apply formulas obtained to study the relative importance of various factors in alveolar carbon dioxide tension and output.

II. *Analysis of CO_2 Flow in Model (Volume Constant)*. Figure 1 indicates schematically the physical relationships which are considered to exist in the alveoli and associated capillaries. This model accounts for all gaseous space in the lungs. Actually, a simpler model may be used for analysis. The space marked Z (since it is neither ventilated nor perfused with capillaries) makes no contribution to the problem; space Y is small and serves only to cut down slightly the effective blood volume flow. Space U is ignored at present for the sake of simplicity but will be included later in an extension of the methods derived by the simpler model. Thus, the model shown in Figure 2 will be used for the basic analysis.

In obtaining a suitable basis for analysis, the following additional assumptions are deemed desirable:

(A) The compartmentation of lung volume into alveoli may be ignored. The CO_2 tension used represents the average value for all alveoli into which there is an active CO_2 flow.

(B) The length of blood flow path is considered to be the same for each blood volume element.

(C) Blood is considered a homogeneous fluid with respect to its ability to hold CO_2 both physically and chemically.

(D) The total cross-sectional area of the capillary path is considered constant throughout the active portion of the path. (This assumption is convenient, though not strictly necessary.)

(E) The capillary channel (actually consisting of a very fine network) is of such a shape that CO_2 diffusion lengthwise along the path is slow compared to CO_2 diffusion out into the alveolar space and therefore may be disregarded.

(F) The individual alveoli are so small and the walls so interlaced with

capillaries that no significant CO_2 tension gradients exist within the alveolar space at any instant of time. At least, any difference may adjust itself in an insignificant amount of time.

(G) Variations in the rate of blood flow are small and are rapid enough so that a constant mean value may be assumed.

(H) Gaseous interchange between the alveoli and the "dead space" is negligible, except during inspiration and expiration. That is, free gaseous diffusion between dead and active space is inconsequential.

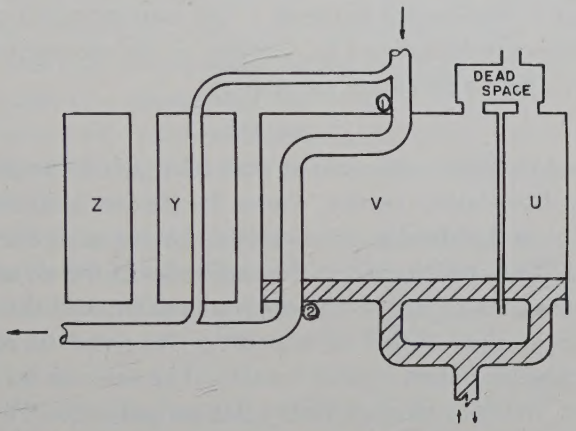

FIGURE 1. Schematic diagram of the physical relationships in the lungs. Space U represents that portion of the lungs which is ventilated but not perfused with capillaries; V is that portion which is both ventilated and perfused; Y is perfused but not ventilated; Z is neither perfused nor ventilated. The process of breathing is depicted as a movement of a piston (hatched) within a cylinder. "Dead space" indicated represents physiological dead space.

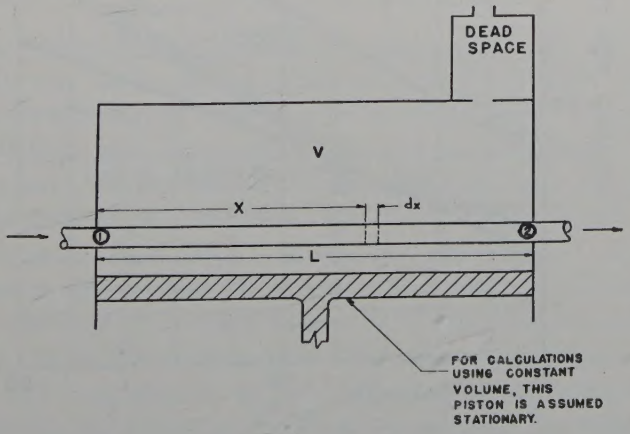

FIGURE 2. Simplified diagram of physical relationships

(I) Temperature and pressure variations are negligible.

Other assumptions will be stated or implied as required in the progress of the analysis.

Let us now define p as the partial pressure of CO_2 in the alveolar space having volume, V. Let M be the molar quantity of CO_2 in space V. The gas constant is R and T is the absolute temperature. Then from the ideal gas law it may be stated that

$$p\,V = MRT\,. \tag{1}$$

We also need a relationship between c, the concentration of CO_2 in the blood, and the corresponding CO_2 tension, q. We assume one of the following form, where a and b are constants:

$$c = aq + b\,. \tag{2}$$

Under general conditions we cannot state that c is a function of q only, for the CO_2 dissociation curves, shown in Figure 3, indicate that the oxygen tension of the blood is also a variable having some effect on c. However, the conditions which exist in the capillaries of the alveoli are reasonably consistent with regard to the oxygen tension, and this fact enabled early workers to draw a line from point ① to point ② representing a "physiological dissociation" curve for CO_2. The exact shape of this curve is not known, but it is most probably flat as indicated. The assumption

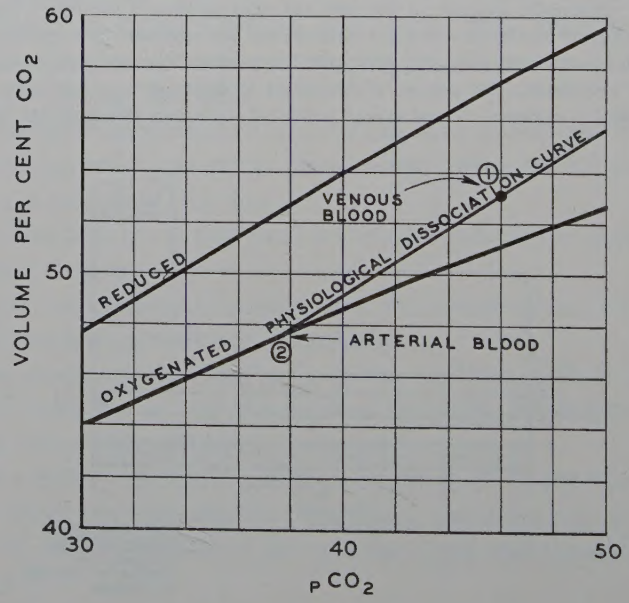

FIGURE 3. A portion of the CO_2 dissociation curve (modified from Peters and Van Slyke)

made herein is that the physiological dissociation curve is close enough to a straight line to enable us to choose a relationship in the form of equation (2) without introducing an appreciable error in results. This approach is not as logical as may be desired, but it provides a convenient way to get away from the only alternate—determination of the complex inter-relationships between chemical-kinetic and physical factors involved.

In solving the problem with constant volume we must define the following quantities: v is the velocity of blood flow through the capillary channel; L is the length of the channel; and B is the cross-sectional area of the channel. Subscripts $_1$ and $_2$ refer to the conditions of the blood at the entrance and exit, respectively, of the capillary channel; $\bar{c}$ refers to the average value of c throughout the capillary channel at any instant of time. The subscript $_0$ refers to known initial conditions, when time (t) is zero. It is apparent that c and q are functions of x and t; whereas p and M are functions of t alone.

Physical conditions prescribe that $p < q$, and therefore CO_2 is passing from the blood to the alveolar space at all times. We see that the rate of increase of the amount of CO_2 in space V is equal to the rate at which CO_2 enters the system at point ① minus the rate at which it leaves the system at point ②, and also minus the rate at which the CO_2 content of the alveolar blood is itself being built up. This gives us then:

$$\frac{dM}{dt} = B\,v\,c_1 - B\,v\,c_2 - BL\,\frac{d\,\bar{c}}{dt}; \tag{3}$$

where

$$\bar{c} = \frac{1}{L}\int_0^L c \cdot dx\,. \tag{4}$$

To obtain a general analytical expression for c, and thence for $\bar{c}$, is quite difficult. We may make an approximation, however, which can be expected to give a negligible error in most practical cases. It is agreed by physiologists that, as a blood volume element passes through the alveolar capillary path, oxygen concentration comes practically into equilibrium with the alveolar pO_2. It is also known that carbon dioxide perfuses tissue much faster than oxygen. Hence, we can normally expect that CO_2 tensions in the alveolar blood and the alveolar gas exist almost in equilibrium for most of the path length, L. The relationship is indicated by Figure 4. (Note: The authors have carried out a more thorough analysis of how CO_2 concentration in the capillaries may be expected to vary with respect to x and t; but, due to limitation of space, the more qualitative reasoning above is considered sufficient for the purpose of this paper.)

We may say then:

$$c \doteq c_2 ; \tag{5}$$

$$q_2 \doteq p . \tag{6}$$

Using equations (1), (2), (5), and (6) to evaluate $\bar{c}$ as defined in equation (4), we find that:

$$\bar{c} \doteq \frac{aRT}{V} M + b . \tag{7}$$

This expression is substituted into equation (3). It is shown in section III that the last term in equation (3), which contains $\bar{c}$, contributes to dM/dt

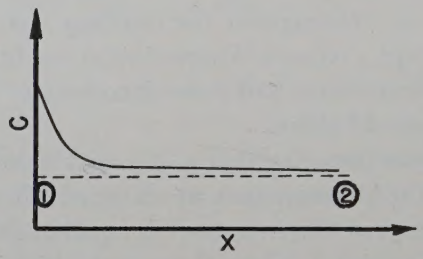

FIGURE 4. Variation of CO_2 concentration of blood in alveolar capillaries at any given instant of time.

only to a minor extent (on the order of 12 per cent); and this fact also minimizes the effect of the error obtained by use of the approximate solution for $\bar{c}$ above. We obtain then:

$$\frac{dM}{dt} = \frac{B V v (c_1 - b)}{V + BLRT a} - \frac{BaRT v}{V + BLRT a} \cdot M . \tag{8}$$

Let:

$$\nu = \frac{1}{\frac{V}{RT a (B v)} + \frac{L}{v}} . \tag{9}$$

Then the solution of the differential equation (8) may be found to be:

$$M = M_0 e^{-\nu t} + \frac{(c_1 - b) V}{aRT} (1 - e^{-\nu t}) . \tag{10}$$

Again using equations (1) and (2), we may transform the above equation to:

$$p = q_1 - e^{-\nu t} \cdot (q_1 - p_0) . \tag{11}$$

III. *Matters in Clarification.* Before proceeding from a discussion of results under constant volume conditions to those under conditions of variable volume, it would be well to observe certain facts which bear on both

aspects. It may be noted that assumptions and approximations have been made on the basis of the numerical size of certain physiological constants and variables. Let us, therefore, list some numerical data (found in the literature) obtained from experimental and theoretical considerations. In attempting to indicate what is "normal" we appreciate the fact that a wide variation in individual values is likely, due to differences among experimental subjects and in experimental techniques. The data included herein are selected so as to use a reasonable average, give greater weight to more recent findings, and "round off" figures slightly for simplicity in calculation.

(A) L/v represents the length of time required for an element of blood to traverse the alveolar capillary. The analyses of F. J. W. Roughton (1945) indicate that this value under resting conditions is about 0.7 seconds. For exercise conditions we feel that 0.4 seconds is approximately correct.

(B) Let us consider total lung volume at inspiration as 3440 cc. and at expiration as 2890 cc., giving a tidal volume of 550 cc. (at body temperature and 100 per cent humidity). If we take dead space volume as 165 cc., then after inspiration V equals 3275 and after expiration V equals 2725. This gives a median value of V as 3000 cc. In exercise we will assume V varies from 3500 cc. to 2000 cc., with a dead space of 250 cc.

(C) From Figure 3 we may arrive at the constants for equation (2). A straight line between points ① and ② is represented by the formula

$$\text{Vol.}\% = 20.8 + \tfrac{5}{7}\text{pCO}_2 . \qquad (12)$$

In the units useful herein, a is easily computed to be about 32×10^{-8} moles/cc. mm. Hg.

(D) Under normal, steady-state conditions the pCO_2 of blood entering the alveolar capillaries is a constant, so that we take q_1 as equal to 44 mm. Hg.

(E) Body temperature used is 310° Kelvin.

(F) In units used herein, the ideal gas constant R is 62,400 mm. Hg. — cc./moles — ° K.

(G) Assuming that almost all the cardiac output passes through the capillaries of the active alveoli, we can see that Bv equals about 5000/60, or 83.3 cc./sec. In exercise, we assume 250 cc./sec.

(H) Roughton's value for the volume of blood in capillaries at any one time (at rest) is around 60 cc. This represents BL. Note that BL equals $Bv \cdot (L/v)$, and this relationship checks very well with data taken above.

Using the above data we can evaluate equation (9) to find that at rest:

$$\nu = \frac{1}{\frac{V}{516} + 0.7} \text{ (sec}^{-1}\text{)} . \tag{13}$$

We are now in position to introduce and utilize a concept called herein "effective volume." It has been noted that in the system represented by Figure (2) the CO_2 is contained not only in the alveolar volume (ignore dead space in this regard) but also within the blood of the alveolar capillaries. As has also been noted, an element of blood during most of its travel down the capillary channel has a CO_2 tension practically in equilibrium with pCO_2 of the alveolar air (see Fig. 4). Hence, any change in alveolar pCO_2, as long as it does not occur too rapidly, has concomitantly a similar change in alveolar blood tension.

In certain calculations it will be found that matters will be simplified if we replace capillary blood by an equivalent volume of gas, representative of the blood in its response to pCO_2 changes. The total of actual alveolar volume plus this hypothetical volume is the effective volume, and may be evaluated as follows:

$$\text{Moles of } CO_2 \text{ in alveolar capillaries} = cBL . \tag{14}$$

Since b moles may be considered mathematically as if permanently bound, therefore:

$$\begin{aligned} \text{Available moles of } CO_2 \text{ in alv. cap.} &= (c - b) \cdot BL \\ &\doteq a p BL . \end{aligned} \tag{15}$$

If this CO_2 were in a gaseous phase, equation (1) would apply and the product of p and hypothetical volume would be equal to

$$(a p BL) \cdot RT . \tag{16}$$

Therefore, the hypothetical volume is equal to

$$aBL \cdot RT . \tag{17}$$

Then, calling the "effective volume" V', we obtain:

$$V' = V + aBLRT . \tag{18}$$

For the normal individual at rest the above equation may be evaluated by substituting numerical data given previously, in which case we have

$$V' = V + 371 \tag{19}$$

with V being around 3000, or about 8 times as large as the hypothetical volume. This enables us to justify the statement made following equation

(7). If the last term of equation (3) had been ignored, we would in effect have substituted the "effective volume" for the actual gas volume in our calculations—an error of about 12 per cent.

IV. *Analysis of CO_2 Flow in Model (Variable Volume).* Having made an analysis under conditions of constant volume, we may now use the results obtained in order to generalize to conditions where the alveolar volume, V, varies with time.

(A) *Variable Volume without Dilution, as during Expiration.* Let us first reduce equation (11) to differential form, and by taking cognizance of its physical significance we can generalize from the constant-volume case, for which equation (11) is valid, to the case of variable volume. If we let

$$y = q_1 - p\,, \tag{20}$$

and put the equation in differential form, we have

$$\frac{dy}{dt} = -\nu y\,. \tag{21}$$

This indicates that whenever a lack of equilibrium exists between CO_2 tensions of a blood volume element and alveolar air, there is a tendency to reduce this difference which is proportional to said difference. The constant of proportionality, ν, represents the effect of the physical and chemical properties of the system and is a function of V as indicated by equation (9), derived on the basis of V being constant. The generalization arises when we state that the tendency to reduce disequilibrium depends on the instantaneous state of the system, in spite of the fact that alveolar volume may be changing with time. This permits us to use equation (21) in the case of variable volume, provided we express V in the equation as a function of time. (Note: This is not strictly true. Analysis by the authors, not included herein, has indicated that volume changes tend to produce transient effects, which are, however, small and persist only a short time. These are therefore ignored.)

Equation (21) may be solved, then, with the knowledge that ν is a function of time. The solution is easily seen to be:

$$y = C \cdot e^{-\int \nu \cdot dt}\,. \tag{22}$$

If V is a linear function of time (we shall so consider it in our analysis), then equation (9) can be written thus:

$$\nu = \frac{1}{s + wt}\,, \tag{23}$$

where

$$s = \frac{L}{v} + \frac{V_0}{RTaBv}, \tag{24}$$

$$w = \frac{\frac{dV}{dt}}{RTaB}. \tag{25}$$

Introducing this relation into equation (22) and solving for y, we obtain:

$$y = \frac{C}{(s + wt)^{1/w}}. \tag{26}$$

If we apply the initial condition that $y = y_0$, when $t = 0$, we arrive at the equation:

$$y = y_0 \left(\frac{s}{s + wt}\right)^{1/w}. \tag{27}$$

Introducing (27) into (20) and putting $y_0 = q_1 - p_0$ we obtain:

$$p = q_1 - (q_1 - p_0)\left(\frac{s}{s + wt}\right)^{1/w}. \tag{28}$$

(B) *Variable Volume with Dilution, as during Inspiration.* Inspiration is a more complex case because, in addition to the effects previously noted, dilution with outside air (so-called tracheal air) or dead space air must be taken into account. For this analysis, assume: (1) that pCO_2 of the diluting gas (designated as p') is constant; and that (2) dilution occurs slowly enough so that CO_2 tension throughout the alveolar system is essentially in equilibrium. The second assumption means that the capillary blood, because of rapid diffusion of CO_2 through the capillary walls, feels the effect of dilution as much as does the alveolar air. We therefore use the "effective volume" concept.

If to the volume V', in which pCO_2 is equal to p, we add a volume ΔV of diluting gas, in which pCO_2 is equal to p', then we have increased the resulting pCO_2 of the system by an amount equal to Δp. By a simple dilution formula we have

$$p + \Delta p = \frac{pV' + p'\Delta V}{V' + \Delta V}. \tag{29}$$

Then,

$$\frac{\Delta p}{\Delta V} = \frac{p' - p}{V' + \Delta V}. \tag{30}$$

Taking ΔV to its zero limit we obtain:

$$\frac{dp}{dV} = \frac{p' - p}{V'}. \tag{31}$$

Since V' is a linear function of t, it may be expressed as:

$$V' = g + ht\,, \tag{32}$$

where

$$g = V_0 + aBLRT, \tag{33}$$

$$h = \frac{dV}{dt}. \tag{34}$$

Then we have

$$dV = dV' = h \cdot dt\,; \tag{35}$$

therefore,

$$\frac{dp}{dt} = \frac{p' - p}{\frac{g}{h} + t}. \tag{36}$$

Adding together the total effects of dilution and CO_2 passage from the blood, expressed by equations (36) and (21), respectively, we obtain the following equation:

$$\left.\begin{aligned} \frac{dp}{dt} &= \frac{p' - p}{\frac{g}{h} + t} + \nu\,(q_1 - p) \\ &= \frac{p' - p}{\frac{g}{h} + t} + \frac{q_1 - p}{s + wt} \end{aligned}\right\}. \tag{37}$$

The above equation is made easy to solve because the denominators of the fractions on the right-hand side are in a simple ratio relation. That is, we can introduce a constant k so that

$$\frac{g}{h} + t = k\,(s + wt)\,. \tag{38}$$

In this equation k is defined either as $1/w$ or as g/hs. By substitution of the more fundamental symbols it is easily shown that from either definition

$$k = \frac{RTaBv}{\frac{dV}{dt}}. \tag{39}$$

We may say then

$$\frac{dp}{dt} = \frac{p' - p}{k\,(s + wt)} + \frac{q_1 - p}{s + wt}. \tag{40}$$

The variables are easily separated, and the solution is found to be:

$$\left[\frac{p'}{k} + q_1 - p \cdot \left(1 + \frac{1}{k}\right)\right]^{-k/k+1} = C \cdot (s + wt)^{1/w}. \tag{41}$$

The initial condition to be taken is that $p = p_0$ when $t = 0$, so that

$$C = \frac{\left(\frac{p'}{k} + q_1 - p_0 - \frac{p_0}{k}\right)^{-k/k+1}}{s^{1/w}}. \tag{42}$$

Introducing (42) into (41) and noting that $k \cdot w = 1$, we obtain:

$$p = \frac{(p' + k\,q_1) - [\,(p' + k\,q_1) - p_0\,(k+1)\,] \cdot \left(1 + \frac{wt}{s}\right)^{-(k+1)}}{k+1}. \tag{43}$$

As a special case, if $p' = 0$ (as when CO_2-free air is inspired),

$$p = \frac{k\,q_1 - [\,k\,q_1 - p_0\,(k+1)\,] \cdot \left(1 + \frac{wt}{s}\right)^{-(k+1)}}{k+1}. \tag{44}$$

V. *Applications of Derived Formulas to Respiratory Problems.* The formulas given by equations (9), (11), (28), (43), and (44) provide the means of computing the changing alveolar pCO_2 and the cyclic CO_2 output of the lungs under equilibrium conditions. Let us apply them to the "normal" data which have been previously listed. Figure 5 shows the respira-

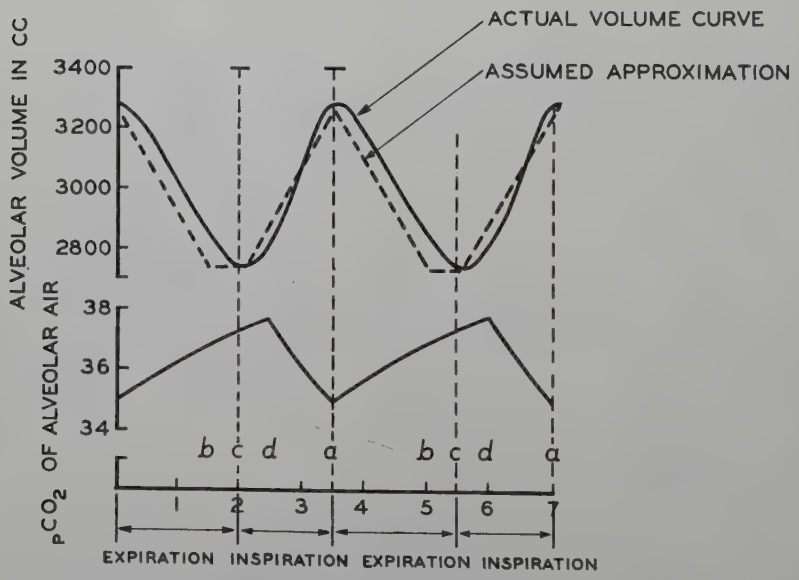

FIGURE 5. Cyclic variation of alveolar pCO_2 and alveolar volume in "normal" respiration. Some pause after expiration has been assumed in the calculations, so that the figure does not indicate the actual agreement between the actual volume curve and the assumed approximate volume curve.

tory pattern which is consistent with these data. For computational purposes, the breathing pattern is approximated rather closely by a series of straight lines. The computations are made separately for each section of the pattern, which can be called phases *a–b*, etc., to correspond to the lettered points on the time axis. Phase *c–d* represents the first portion of inspiration, during which dead space air is being re-inspired. Point *d* therefore represents the instant when fresh air from without is first drawn into the alveoli.

In phase *a–b*, the expiration phase,

$$V = 3275 - 367t\,, \tag{45}$$

and, by use of equations (9) and (28), we can find

$$p_b = 9.02 + 0.795p_a\,. \tag{46}$$

It is necessary to compute the average value of p for the last 0.45 seconds of phase *a–b*. This is the time necessary for expulsion of 165 cc. of air from the alveoli, which portion does not escape to the outside atmosphere but remains in the dead space to be mixed and re-introduced to the alveolar space during the first 0.45 seconds of the next inspiration. This average is obtained by integrating equation (28) over the last 0.45 seconds of the expiration and dividing by the time increment. This will give the following value of p' for the phase *c–d:*

$$p'_{cd} = 7.79 + 0.823p_a\,. \tag{47}$$

Phase *b–c* is the resting phase, at constant volume, and is computed simply by use of equations (9) and (11). We obtain:

$$p_c = 3.52 + 0.920p_b\,. \tag{48}$$

In phase *c–d*, the first part of inspiration,

$$V = 2725 + 367t\,, \tag{49}$$

and, by use of equations (9), (43), and (47), it is found that

$$p_d = 3.45 + 0.0406p_a + 0.881p_c\,. \tag{50}$$

In phase *d–a*, the latter part of inspiration,

$$V = 2890 + 367t\,, \tag{51}$$

and, by use of equations (9) and (44), one determines that:

$$p_a = 6.06 + 0.764p_d\,. \tag{52}$$

Equations (46), (48), (50), and (52) constitute a set of four equations

in four unknowns which can be solved simultaneously to give us:

$$\left.\begin{aligned} p_a &= 34.90 \\ p_b &= 36.74 \\ p_c &= 37.32 \\ p_d &= 37.75 \end{aligned}\right\}. \tag{53}$$

These values are plotted in Figure 5, below the respiration pattern curve.

The alveolar air which is expelled and lost to the outer atmosphere is that portion eliminated from the alveolar space during the first 1.05 seconds of phase *a–b*. Let W be the volume of air expelled from the alveolar space from beginning of expiration to any arbitrary time. Then,

$$W = 367t\,. \tag{54}$$

In the expelled air the amount of CO_2 in a volume increment, dW, can be called dM_1.

$$\left.\begin{aligned} dM_1 &= \frac{p}{RT} \cdot dW \\ &= \frac{367p}{RT} \cdot dt \end{aligned}\right\}. \tag{55}$$

By integrating and substituting proper numerical data, we find that

$$\left.\begin{aligned} M &= \frac{1}{1410} \text{ moles/cycle} \\ &= 272 \text{ cc/min(STP)} \end{aligned}\right\}. \tag{56}$$

This figure may be checked by calculating the amount of CO_2 dumped according to the dissociation curve (Fig. 3). From the data calculated we judge that the average value of p is about 36.3. If the blood enters with a pCO_2 of 44 mm. and leaves on the average with a pCO_2 of 36.3 mm., the pCO_2 has decreased 7.7 mm. This represents a difference in volume per cent of 5.5. If 5000 cc. of blood pass through the active alveoli per minute, the amount of CO_2 emitted is $(5000/100) \times 5.5$, or 275 cc./min. This checks quite closely with the result found by the previous method.

It may be noted at this point that the compound curve in Figure 5, from p_a to p_d, resembles a simple exponentially rising curve and suggests that perhaps adequate results could be obtained by use of a single simple formula for that portion rather than several more complex ones developed herein. As a matter of fact, equations (9) and (11) will serve very well provided that a median value of V is used for the determination of the equation of the curve. By employing this method results were obtained as

follows:

$$\left.\begin{aligned} p_a &= 34.90 \text{ (assumed as basis)} \\ p_b &= 36.78 \\ p_c &= 37.30 \\ p_d &= 37.75 \end{aligned}\right\}. \qquad (57)$$

Comparison with the results, listed as equation (53), indicates the accuracy that may be obtained with this simplified method. Such accuracy of course could not be expected if alveolar volume should vary widely over the period of the cycle.

Apparently, the shape of the expiration portion of the respiration volume curve (Fig. 5) makes little difference since the rise in pCO_2 in the alveoli depends primarily upon the median volume in the cycle.

All the previous results and calculations have been based on what is called the "normal" individual's characteristics while at rest. There are wide physiological variations from these characteristics under equilibrium conditions of exercise, of varying external environment, or when pathology exists. It is of interest therefore to repeat the previous calculations for different conditions. Table I indicates the variations in data used and in

TABLE I

THEORETICAL ANALYSIS OF THE INFLUENCE OF CHANGES IN SOME PHYSIOLOGICAL FACTORS ON CO_2 OUTPUT AND ALVEOLAR pCO_2

Case	Change in Physiological Factors		Change in Computed Results		
	Altered Factor	Amount	CO_2 Output	Median Alv. pCO_2 (mm. Hg.)	Cyclic pCO_2 Variation (mm. Hg.)
1...	Slope of CO_2 Diss. Curve	+10%	+ 1%	+0.5	−0.1
2...	Av. Alveolar Volume	− 8%	0	0	+0.2
3...	Venous CO_2 Tension	+1 mm. Hg.	+ 2%	+0.8	+0.1
4...	Total blood flow rate	+10%	+ 1%	+0.5	−0.1
5...	Blood Time in Alv. Cap.	+14%	0	−0.1	−0.1
6...	Breathing rate	+75%	+55%	−4.3	−0.5
7...	Moderate Exercise	.....	+640%	.....	+5.7
8...	Tidal Volume	− 9%	−11%	+0.8	−0.3
9...	Dead Space	− 9%	+ 3½%	−0.3	0
10...	Inact. Part of Alv. Vol.	9%	− 6%	+0.4	0

results obtained. In this table, data are given relative to the basic solution for "normal" conditions. With one exception, as noted below, the calculations follow almost precisely the pattern presented above.

It will be noticed that for Case 10 of Table I, 9 per cent of the alveolar

volume is considered as inactive in the calculations. This means that we have used a model more complicated than that of Figure 2, for we have taken into account both spaces U and V of Figure 1. The procedure hitherto described is complicated by the following facts: (1) during expiration part of the expired air comes from the inactive as well as the active volume; (2) dead-space air which is re-inhaled is made up of air from U as well as from V, in proportion to the relative volumes of these spaces; (3) dilution by inhalation affects the air of both spaces, U and V. In spite of these complications the calculation is straight-forward and need not be explicitly detailed. Results in this case are:

$$\left.\begin{aligned} p_a &= 35.30 \\ p_b &= 37.23 \\ p_c &= 37.85 \\ p_d &= 38.20 \\ f_a &= f_b = f_c = 10.29 \\ f_d &= 11.67 \\ M_1 &= 257 \text{ cc./min.(STP)} \end{aligned}\right\} \qquad (58)$$

(f refers to pCO_2 for space U, as p does for space V).

From the results of the analyses, indicated by Figure 5 and Table I, several important facts may be discerned with regard to external CO_2 respiration.

(A) The result of the basic case—taken as "normality"—gives a value of CO_2 output which is reasonably close to that value generally accepted as average for human beings. The classic value is about 20 per cent lower, but considering the vagueness of some of the basic data, this may be considered good agreement. It might be noted that Case 10 of Table I, which is probably closer to actual physiological conditions than the basic solution, gives a value of CO_2 output which is lower and closer to the traditionally accepted value.

(B) It is interesting to note the relative importance of various physiological factors in their ability to affect directly average pCO_2 of the alveoli, cyclic variation in pCO_2 and total CO_2 output. Of particular interest is the fact that a change in average alveolar volume, or a change in the length of time a blood element takes to pass through the alveolar capillaries, has little effect on any of the results. Also, we note that an increase in CO_2 tension of the incoming blood does raise somewhat the alveolar pCO_2 but has much less effect on CO_2 output than one might expect. (This does not take into account the effect a change in one factor may have on another.

See paragraph below.) On the other hand, an increase in breathing rate has a very definite effect on all results. Exercise involves a radical change in many physiological conditions and has a great effect on alveolar CO_2 respiration.

A warning must be inserted here. A change in one of the physiological factors in a given individual involves more than the direct results indicated above. If the total picture is desired, it is necessary to take into account other physiological adjustments via reflex pathways and chemical effects on the respiratory center. These factors have been discussed by J. S. Gray (1950). For example, should the cardiac output increase 10 per cent, the resulting 0.5 mm. pCO_2 increase of alveolar air (hence, of arterial blood) would cause respiratory center stimulation so that CO_2 output would increase. Future studies which *do* take into account such readjustments among all the respiratory factors are indicated.

(C) On the assumption that blood as it leaves the alveolar capillaries is practically in equilibrium with alveolar pCO_2, we see that the blood normally undergoes a decrease of about 7.5 mm. Hg. in CO_2 tension in its passage through the lungs. Or, if one prefers to consider the model proposed by Case 10, this drop becomes about 7.0 mm. Hg. Such a value is very close to experimentally determined results.

(D) One of the most interesting predictions of these calculations is the cyclic variation in alveolar pCO_2. This fact is one which has recently been verified experimentally, though not yet quantitatively measured in man. Our calculations show a fluctuation of 2.5 to 3.0 mm. Hg. during the respiratory cycle. Actually, the sharp breaks in the pCO_2 curve shown in Figure 5 are most likely rounded off. We can estimate that actual experimental values (should they become available) would be somewhere between 2.0 and 2.5 mm. Hg. Note should be made of the large variation predicted for exercise conditions. This should indicate to workers an experimental condition more amenable to measurement than the normal.

In addition to providing general light on external CO_2 respiration, as indicated above, the results tabulated may be used to solve many concrete problems. In general two kinds of problems may be attacked by use of Table I. The first type is the analysis of respiratory conditions in an individual whose physiological factors in equilibrium differ from those assumed by us as "normal." The second type of problem is the study of the results of factor changes in a single individual. An example of each type of problem follows.

Problem (1). If an individual in a state of metabolic equilibrium exhibits all the characteristics listed herein as "normal," except that because of

individual variation in the ability of blood to carry CO_2 the slope of the dissociation curve is 20 per cent less, what difference from the "normal" might be expected?

Answer. From the tabulated data, the only difference we can see is that the average alveolar pCO_2 would be about 1 mm. Hg. lower than that indicated for Case 1, Table I (normality). This difference assumes that concomitant with the change in slope of the dissociation curve, there is a difference of sensitivity of the respiratory center such that alveolar ventilation is maintained at the level used in the original calculation.

Problem (2). If an individual, after attaining equilibrium at rest, voluntarily hyperventilates by increasing the depth of respiration 18 per cent what is the effect on the results?

Answer. Such an increase in tidal volume is seen to cause an increased CO_2 output of about 22 per cent, along with a decrease of 1.6 mm. Hg. for alveolar pCO_2 and an increased cyclic variation in same. If the individual remained in the original resting condition, physiological changes would occur, and his body would not continue the increased CO_2 output. The nature of the compensatory changes is not predicted herein. However, two changes might be expected: a reduction in venous CO_2 tension and a decreased breathing rate. A drop in CO_2 tension of entering blood of $5\frac{1}{2}$ mm. and a reduction in breathing rate of 15 per cent would approximately restore the respiratory CO_2 output to equilibrium with the metabolic output.

Thanks are due to Professor H. D. Landahl for a suggestion which provided a neat and shortened analysis of the basic problem.

LITERATURE

Best, C. H., and N. B. Taylor. 1945. *The Physiological Basis of Medical Practice.* Baltimore: Williams and Wilkins Co.

Gray, J. S. 1950. *Pulmonary Ventilation and Its Physiological Regulation.* American Lecture Series No. 63. Springfield, Ill.: Charles C. Thomas.

Hitchcock, F. A., and R. W. Stacy. 1950. "A mass spectrometer for study of respiratory gases." Joint AIEE-IRE Conference on Instrumentation in Nucleonics and Medicine, New York, November, 1950.

Hunter, J. A., R. W. Stacy, and F. A. Hitchcock. 1949. "A mass spectrometer for continuous gas analysis." *Rev. Sci. Inst.*, **20**, 5, 333.

Nims, L. F. 1949. "Respiration." (Section VII of Fulton, *Textbook of Physiology.*) Philadelphia: W. B. Saunders Company.

Peters, J. P., and D. D. Van Slyke. 1931. *Quantitative Clinical Chemistry*, Volume I: *Interpretations.* Chapter 12, "Hemoglobin and Oxygen." Baltimore: Williams and Wilkins Company.

Roughton, F. J. W. 1945. "The average time spent by the blood in the human lung capillary and its relation to the rates of CO uptake and elimination in man." *Am. Jour. Physiol.*, **143**, 621.

Stacy, R. W., J. A. Hunter, and F. A. Hitchcock. 1948. "A mass spectrometer for the rapid continuous analysis of respiratory gases." *Fed. Proc.*, **7**, 1.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

ON THE STEADY LAMINAR FLOW OF A VISCOUS INCOMPRESSIBLE FLUID IN AN ELASTIC TUBE

G. W. Morgan
Brown University

The conditions are examined under which an approximate relation between the radius of the tube and the distance along the axis, as obtained by N. Rashevsky (1945), is consistent with the assumptions made in the solution. These conditions are reduced to relations between two dimensionless parameters of the system. Second approximations are found for three different ranges of values of these parameters.

Whenever a viscous fluid flows through a tube, there must be a drop in fluid pressure along the tube to overcome the frictional resistance at the walls. As a result of this pressure gradient and the elasticity of the tube wall, the radius of the tube will decrease with distance along the axis. Hence the boundary condition of vanishing velocity at the wall has to be applied at an unknown boundary.

This problem was studied by N. Rashevsky (1945) who found a solution to a first degree of approximation and then used this solution to obtain a second approximation. We shall here consider the same problem in more detail.

Fundamental Equations. We use a cylindrical coordinate system r, θ, z in which r is the radius and z is the distance along the axis of the tube. The cross-section of the tube is assumed to be circular; hence we assume axial symmetry, so that the velocity component v_θ and its derivatives vanish. We then have the following fundamental equations.

Equations of Motion

i) Equation of continuity:

$$\frac{1}{r}\frac{\partial r v_r}{\partial r}+\frac{\partial v_z}{\partial z}=0\,, \tag{1}$$

where v_r and v_z are the radial and axial velocity components respectively.

We can also write an integral form of the continuity equation, viz., that the total mass flow across any tube section of radius $R(z)$ is constant and

equal to, say, Q:

$$2\pi\rho\int_0^{R(z)} r\, v_z dr = Q. \tag{2}$$

ii) Navier-Stokes equations for axially-symmetric, steady, laminar flow of an incompressible fluid:

$$v_r\frac{\partial v_r}{\partial r}+v_z\frac{\partial v_r}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial r}+\nu\left(\frac{\partial^2 v_r}{\partial r^2}+\frac{1}{r}\frac{\partial v_r}{\partial r}+\frac{\partial^2 v_r}{\partial r^2}-\frac{v_r}{r^2}\right) \tag{3}$$

$$v_r\frac{\partial v_z}{\partial r}+v_z\frac{\partial v_z}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial z}+\nu\left(\frac{\partial^2 v_z}{\partial r^2}+\frac{1}{r}\frac{\partial v_z}{\partial r}+\frac{\partial^2 v_z}{\partial z^2}\right), \tag{4}$$

where p is the pressure (it is convenient to take p as the difference between the fluid pressure and the constant pressure prevailing outside the tube); ρ is the density, a constant; ν is the kinematic viscosity $\equiv \mu/\rho$, where μ is the coefficient of viscosity.

iii) Force per Unit Area. The normal compressive force on a unit area perpendicular to a radius is given by:

$$E_{rr}(r,z) = p - 2\mu\frac{\partial v_r}{\partial r}. \tag{5}$$

Equation of Elasticity

If R varies very slowly with z and if we neglect effects at the ends of the tube, then the forces in the tube will have no component parallel to the axis. Then, if we assume a linear stress-strain relation, the following equation can be derived, provided δ, the wall thickness, is very small compared with R:

$$E_{rr}(R,z) = E\delta\left(\frac{1}{R_u}-\frac{1}{R}\right), \tag{6}$$

where E is Young's modulus, R_u is the unstretched radius of the tube (subsequently assumed to be constant), and E_{rr} represents the difference between the normal compressive stress on the interior of the tube and the external pressure.

It should be noted that in the derivation of this equation we neglect the effect on the radius $R(z)$ of the fluid friction on the interior of the tube, and that, due to the slow variation of $R(z)$, we take the component of E_{rr} which is normal to the wall to be equal to E_{rr} itself.

If the strain $(R - R_u)/R_u$ is so large that the stress-strain relation cannot be assumed to be linear, but the variation in R over the length of tube considered is small, so that the range of strain along the tube is small enough, then (6) can be replaced by

$$E_{rr}(R,z) - p_s = E_s\delta\left(\frac{1}{R_s}-\frac{1}{R(z)}\right), \tag{7}$$

where p_s is the constant difference between internal and external pressure when the tube is stretched uniformly under hydrostatic pressure; p_s has a value close to $E_{rr}(R, z)$ in the actual problem; R_s is the uniform radius of the tube when subjected to p_s; E_s is the slope of the stress-strain curve at the point where the strain is $(R_s - R_u)/R_u$.

Equation (7) is the result of assuming a linear variation of stress with strain over a small range around $(R_s - R_u)/R_u$.

The following treatment uses (6) but is equally valid for (7), provided E_s be substituted for E.

First Approximation to the Solution. If $R(z)$ were a constant then an exact solution of equations (1), (3) and (4) would be the flow commonly known as Poiseuille flow, in which

$$v_r = 0 \qquad \text{and} \qquad v_z = \frac{1}{4\mu}\frac{dp}{dz}(r^2 - R^2).$$

This flow would occur if the tube were rigid; practically, this means that the tube would have to be sufficiently "strong" so that variations in the fluid pressure along the tube would have a negligible effect on the tube radius. We shall see later more precisely what is meant by a "strong" tube in connection with this problem.

Following Rashevsky, we now assume that in the case of a non-rigid tube we can, to a first approximation, neglect all terms containing v_r or its derivatives, i.e., that the tube is strong enough so that $R(z)$, though no longer a constant, varies so slowly that the resulting radial velocity and terms involving it are negligibly small. We shall investigate a posteriori under what conditions this assumption may be expected to yield reasonable results. We now proceed without further assumptions.

From (3) we find: $\partial p/\partial r = 0$, i.e., $p = p(z)$ and hence from (5) it follows that: $E_{rr}(R, z) = p(z)$. Then (6) gives us

$$\frac{dp}{dz} = \frac{E\delta}{R^2}\frac{dR}{dz}, \tag{8}$$

and (4) becomes

$$\frac{1}{\rho}\frac{dp}{dz} = \nu\left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r}\frac{\partial v_z}{\partial r}\right). \tag{9}$$

Integration of (9) and application of the boundary condition $v_z = 0$ at $r = R$ yields:

$$v_z = -\frac{1}{4\mu}\frac{dp}{dz}[R^2(z) - r^2]. \tag{10}$$

For any fixed z, (10) gives a parabolic radial distribution for v_z. Hence

we see that if we neglect terms containing v_r, then the velocity profile at each section must necessarily be the same as in Poiseuille flow.

Substitution of (8) into (10) gives us:

$$v_z = -\frac{E\delta}{4\mu}\frac{dR}{dz}\left(1-\frac{r^2}{R^2}\right). \tag{11}$$

Finally we apply (2), the continuity condition, and obtain:

$$\frac{dR}{dz} = -\frac{8\mu Q}{E\delta\pi\rho R^2} \tag{12}$$

and

$$\frac{R^3}{3} = \frac{R_0^3}{3} - \frac{8\mu Q}{E\delta\pi\rho}z, \tag{13}$$

where R_0 is the radius at $z = 0$.

Equation (13) is identical with Rashevsky's result. The method of its derivation, however, differs slightly from Rashevsky's in that the parabolic variation of v_z with radius [equation (10)] was shown to be a consequence of the assumption concerning the smallness of v_r and its derivatives, rather than an additional assumption.

Validity of the Approximation. We can now examine under what circumstances the above solution will be consistent with the assumptions made provisionally in the section above.

The exact equation (4) may be replaced by the approximate equation (9) provided

$$v_r\frac{\partial v_z}{\partial r} + v_z\frac{\partial v_z}{\partial z} \quad \text{and} \quad \nu\frac{\partial^2 v_z}{\partial z^2}$$

are very small compared with

$$\nu\left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r}\frac{\partial v_z}{\partial r}\right).$$

From (8), (9), and (12) we have

$$\nu\left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r}\frac{\partial v_z}{\partial r}\right) = -\frac{8\mu Q}{\pi\rho^2 R^4}.$$

Also, using our results, the following orders of magnitude are easily verified:

$$\nu\frac{\partial^2 v_z}{\partial z^2} = O\left(\frac{\mu^3 Q^3}{E^2\delta^2\rho^4 R^8}\right),$$

$$v_r\frac{\partial v_z}{\partial r} = O\left(\frac{\mu Q^3}{E\delta\rho^3 R^7}\right),$$

$$v\frac{\partial v_z}{\partial z} = O\left(\frac{\mu Q^3}{E\delta\rho^3 R^7}\right).$$

Hence we may neglect the non-linear terms, provided

$$\frac{Q^2}{E\delta\rho R^3} \ll 1 . \tag{14}$$

Or, writing $Q = \rho V R^2$ where V is an average velocity over a cross section, we have

$$\frac{\rho V^2}{\frac{E\delta}{R}} \ll 1 . \tag{15}$$

We may also neglect the term $\nu(\partial^2 v_z/\partial z^2)$, provided

$$\frac{Q^2\mu^2}{E^2\delta^2\rho^2 R^4} \ll 1 \tag{16}$$

or

$$\left(\frac{\frac{\mu V}{R}}{\frac{E\delta}{R}}\right)^2 \ll 1 . \tag{17}$$

If we write $\rho V R/\mu = \mathfrak{R}$, a Reynolds number appropriate to our problem (V and R may be taken as the values at $z = 0$) and $\mu V/E\delta = \mathfrak{N}$, then conditions (15) and (17) become respectively

$$\mathfrak{R}\mathfrak{N} \ll 1 \tag{18}$$

and

$$\mathfrak{N}^2 \ll 1 . \tag{19}$$

We note that $\mathfrak{N}$ is a dimensionless ratio of a viscous to an elastic force. We may regard $\mathfrak{R}$ and $\mathfrak{N}$ as the two dimensionless parameters of the system. For any fixed $\mathfrak{R}$ the flow will approach Poiseuille flow as $\mathfrak{N} \to 0$. Our provisional assumptions also involved equations (3) and (5). We can write $\partial E_{rr}/\partial z = \partial p/\partial z$ provided $|\mu(\partial^2 v_r/\partial r \partial z)| \ll |\partial p/\partial z|$ and this, using our results and equation (1), is found to be true if $\mathfrak{N}^2 \ll 1$.

Finally, using (3), we find

$$\frac{1}{\rho}\frac{\partial p}{\partial r} = -\frac{2\nu Q}{\pi\rho}\frac{dR}{dz}\left(\frac{8r}{R^5}\right) + \text{higher order terms.}$$

By integrating from r to R and using (6) we obtain

$$p(r,z) = E\delta\left(\frac{1}{R_u} - \frac{1}{R}\right) + \frac{8\mu Q}{\pi\rho}\frac{dR}{dz}\left(\frac{1}{R^3} - \frac{r^2}{R^5}\right) \tag{20}$$

whence

$$\frac{\partial p}{\partial z} = \frac{E\delta}{R^2}\frac{dR}{dz} + \frac{8\mu Q}{\pi\rho}\left(\frac{dR}{dz}\right)^2\left(-\frac{3}{R^4} + \frac{5r^2}{R^6}\right) + \frac{8\mu Q}{\pi\rho}\frac{d^2R}{dz^2}\left(\frac{1}{R^3} - \frac{r^2}{R^5}\right). \tag{21}$$

The last two terms are negligible, provided $\mu^2 Q^2/E^2\delta^2\rho^2 R^4 \ll 1$ or $\mathfrak{N}^2 \ll 1$. Thus (18) and (19) are the conditions which must be satisfied.

Let us first consider the case of $\mathfrak{R} \gg 1$. If (18) is satisfied then (19) will be automatically satisfied, since it is a weaker condition in the present case. Hence, for $\mathfrak{R} \gg 1$, the results are consistent with our assumptions if $\mathfrak{N}$ is so small that $\mathfrak{R}\mathfrak{N} \ll 1$.

If $\mathfrak{R} = O(1)$ then (18) is true if $\mathfrak{N} \ll 1$. Hence in this case this is the only condition.

If $\mathfrak{R} \ll 1$ and $R = O(\mathfrak{N})$ the results and assumptions are consistent if $\mathfrak{N}^2 \ll 1$.

We now see that a "strong" or "practically rigid" tube in connection with this problem means that $\mathfrak{R}\mathfrak{N}$ and $\mathfrak{N}^2$ are very much less than one.

Second Approximation. Cases $\mathfrak{R} \gg 1$, $\mathfrak{R}\mathfrak{N} \ll 1$ and $\mathfrak{R} = O(1)$, $\mathfrak{N} \ll 1$. We can obtain better results than the foregoing by taking account of some of the terms previously neglected by a method of successive approximations: we shall include terms of second order by computing their approximate value from the first approximation. Calling the terms retained in the first approximation terms of order one, the next higher order terms in the present case will be those of order $\mathfrak{R}\mathfrak{N}$, i.e., the quadratic terms in (4). All other terms are of order $\mathfrak{N}^2$ which is negligible compared with $\mathfrak{R}\mathfrak{N}$. Hence we still have $\partial p/\partial r = 0$ and (8) holds. It is important to note that the quadratic terms $v_r(\partial v_z/\partial r)$ and $v_z(\partial v_z/\partial z)$ are both of order $\mathfrak{R}\mathfrak{N}$ and that both must therefore be included in the present approximation.

Using the results of the second section and equation (1) these terms can be written as follows:

$$v_r \frac{\partial v_z}{\partial r} + v_z \frac{\partial v_z}{\partial z} \equiv I(r,z) = \frac{8Q^2}{\pi^2\rho^2} \frac{dR}{dz}\left(-\frac{r^4}{R^9} + \frac{2r^2}{R^7} - \frac{1}{R^5}\right). \tag{22}$$

Hence (4) in our second approximation becomes:

$$I(r,z) = -\frac{1}{\rho}\frac{dp}{dz} + \nu\left(\frac{\partial^2 v_z}{\partial r^2} + \frac{1}{r}\frac{\partial v_z}{\partial r}\right). \tag{23}$$

Our first solution resulted in a parabolic radial distribution of v_z. This was essentially due to the fact that in the approximate equation (9) the viscous term was not a function of r. In (23) this is no longer the case and hence we expect a modified velocity profile. Integrating (23), using (8), and applying the boundary condition $v_z = 0$ at $r = R$, we obtain

$$v_z = -\frac{E\delta}{4\mu}\frac{dR}{dz}\left(1 - \frac{r^2}{R^2}\right) + \frac{2Q^2}{\pi^2\rho\mu}\frac{dR}{dz}\left(-\frac{r^6}{9R^9} + \frac{r^4}{2R^7} - \frac{r^2}{R^5} + \frac{11}{18R^3}\right). \tag{24}$$

Comparing equation (24) with (11) we see that the second term in (24)

is the additional term introduced by taking account of the inertia terms. Finally, applying (2), the following equation is obtained for dR/dz:

$$\left(-\frac{E\delta\pi\rho}{8\mu Q}R^2+\frac{Q}{2\pi\mu}\frac{1}{R}\right)\frac{dR}{dz}=1\,. \qquad (25)$$

Since dR/dz is negative we see that the effect of the correction term $Q/2\pi\mu R$ [compared with the corresponding term of (12)] is to cause R to decrease more rapidly. Equation (24) shows that for a given $R(z)$ the correction term tends to reduce the speed on the axis and to maintain the given volume flow by a sharper gradient near the wall.

We note that (24) and (25) hold subject to the previously discussed restrictions; moreover, since the conditions (14) and (16) involve the variable R, the equations only hold over a range of z in which R does not become too small.

The above case is the one treated by Rashevsky, who obtains an equation identical with our (25) except for the fact that the numerical coefficient $\frac{2}{3}$ of his correction term is replaced in (25) by $\frac{1}{2}$. This discrepancy is due to the fact that Rashevsky only takes account of the second of the non-linear terms in (4) and, moreover, that he applies some average value of this term over the cross-section. This latter step results in the viscous term of (23) being independent of r as it is in the first approximation and hence in the retention of the parabolic velocity profile.

Second Approximation. Case $\mathfrak{R} \ll 1$. Here the first approximation would hold even if $\mathfrak{N} = O(\mathfrak{R})$. The terms to be taken into account in the second approximation will depend on the order of $\mathfrak{N}$. If $\mathfrak{N} \ll \mathfrak{R}$, then a second approximation will hardly be useful since the correction terms will be very small. If $\mathfrak{N} = O(\mathfrak{R})$ then, in the second approximation, we must take account of terms of order $\mathfrak{R}\mathfrak{N}$ and $\mathfrak{N}^2$, i.e., of equation (3) as well as all terms previously neglected in (4) and (5).

This can be done by a method analogous to that used in the preceding section by replacing all unknown terms of second order by their approximate values as obtained from the first approximation.

Using this process, equation (4) becomes

$$I(r,z) = -\frac{1}{\rho}\frac{\partial p}{\partial z}+\nu\left(\frac{\partial^2 v_z}{\partial r^2}+\frac{1}{r}\frac{\partial v_z}{\partial r}\right)+\nu G(r,z)\,, \qquad (26)$$

where $G(r, z)$ is the known first approximation of $\partial^2 v_z/\partial z^2$. Also we have

$$\frac{\partial p}{\partial z}=\frac{E\delta}{R^2}\frac{dR}{dz}+H(r,z)\,, \qquad (27)$$

where $H(r, z)$ is a known function which corrects for the fact that $\partial p/\partial r$

is no longer negligible. The function H is determined from equations (3), (5), and (6).

If we introduce expression (27) for $\partial p/\partial z$ into (26), integrate with respect to r and apply the boundary condition on $r = R$, then application of (2) will yield the new equation for dR/dz. It is:

$$\left(-\frac{E\delta\pi\rho}{8\mu Q}R^2+\frac{52}{3}\frac{\mu Q}{E\delta\pi\rho}\frac{1}{R^2}+\frac{Q}{2\pi\mu}\frac{1}{R}\right)\frac{dR}{dz}=1\,. \tag{28}$$

Because of the very small values of $\mathfrak{N}$ which are likely to be encountered in practice, it is to be expected that for most applications $\mathfrak{N} \ll \Re$; in this case the second term in parentheses in (28) becomes negligible compared with the third term and the equation reduces to (25).

LITERATURE

Rashevsky, N. 1945. "A Problem in Mathematical Biophysics of Blood Circulation: II." *Bull. Math. Biophysics*, **7**, 35–9.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

THE AUDIOGYRAL ILLUSION AND THE MECHANISM OF SPATIAL REPRESENTATION

ROBERT MAYNE
GOODYEAR AIRCRAFT CORPORATION
AKRON, OHIO

An hypothesis regarding the mechanism of spatial representation in the neural centers is formulated in order to explain the audiogyral illusion. Using this hypothesis and experimental data (Clark and Graybiel, 1949) the time constant of the semicircular canals is calculated. The value obtained in this manner agrees with that previously calculated (Mayne, 1950) using results from experiments on the nystagmic latency of pigeons.

The Audiogyral Illusion. Clark and Graybiel (*loc. cit.*) determined the effect of angular rotation on the localization of sounds using human subjects. A mistaken judgment regarding the direction of sound was noted during the period immediately following rotation. The phenomenon was called an *audiogyral illusion.*

The paper by Clark and Graybiel gives a complete description of the experimental procedure. In brief, however, the subjects were first rotated at a constant velocity in a Barany chair, then stopped suddenly. They were then asked to judge the apparent direction of a sound source located at a fixed position with respect to themselves and to the chair. The apparent displacement of the source from its actual position was noted as a function of the elapsed time after rotation was stopped. A total of six series of tests was made, for clockwise and counterclockwise rotation and for three initial positions of the sound with respect to the observer—ten degrees to the right, ten degrees to the left, and center.

A number of control runs were also made to evaluate the ability of the subjects to estimate correctly the direction of the sound under static conditions, i.e., without previous rotation. It was found that the mean error for the three positions was only two degrees.

The results of the test are best summarized in the words of the original paper:

> Clearly defined changes in sound localization occurred following the angular decelerations used in this experiment. The changes were consistent for the three subjects and showed that after abrupt decelerations, they tended to localize the

sound as if it were displaced in the direction of the preceding rotation. Thus, after rotation to the left, a sound coming from a source directly in front of the S's would be reported as coming from the left of center. On the other hand the illusory feelings of rotation would be opposite in direction, i.e., the S would report that he felt as if he were rotating to the right.

An attempt will now be made to interpret the results and to fit these results into a theoretical framework of the behavior of the semicircular canals and of the neural representation of space.

The Post Rotational Effects. Since the audiogyral illusion takes place during the period immediately following the cessation of a steady rotation,

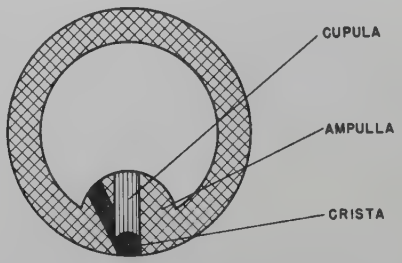

FIGURE 1. Schematic representation of semicircular canal (*Jour. Comp. and Physiol. Psychol.*, **43**, No. 4, August, 1950).

we should consider briefly the behavior of the semicircular canals during this particular period.

In a previous paper (Mayne, 1950) the thesis that the semicircular canals issue signals which are proportional to the angular velocity of the head was developed. It was shown that this proportionality obtains only within certain ranges of amplitude and frequencies corresponding to normal body movements. During the post rotational period the semicircular canals issue definite velocity signals although the head is stationary. A quantitative prediction of these post rotational signals can be made with a brief consideration of the mechanism of the semicircular canals.

Referring to Figure 1, illustrating diagrammatically the structure of the semicircular canals, the following equation can be seen to describe the motion of the endolymph:

$$m\left(\frac{d^2\theta_\omega}{dt^2}-\frac{d^2\theta}{dt^2}\right)=c\frac{d\theta}{dt}+K\theta. \qquad (1)$$

This may be written as:

$$\frac{d^2\theta_\omega}{dt^2}=\frac{d^2\theta}{dt^2}+\frac{c}{m}\frac{d\theta}{dt}+\frac{K\theta}{m}, \qquad (2)$$

where:

m = mass of the system

c = viscosity of the system

K = elasticity of the system

θ_ω = displacement of the walls of the canals in space

θ = displacement of the fluid with respect to the walls.

From the previous paper (Mayne, 1950) we have:

$$\frac{c}{m} = 200\,,$$

$$\frac{K}{m} = 24\,.$$

Equation (2) then becomes:

$$\frac{d^2\theta_\omega}{dt^2} = \frac{d^2\theta}{dt^2} + 200\,\frac{d\theta}{dt} + 24\theta\,. \tag{3}$$

Figure 2 illustrates the displacement of the fluid, as calculated from equation (3), when the body is accelerated abruptly to a constant velocity; then, following a period of half a minute or more, is suddenly stopped.

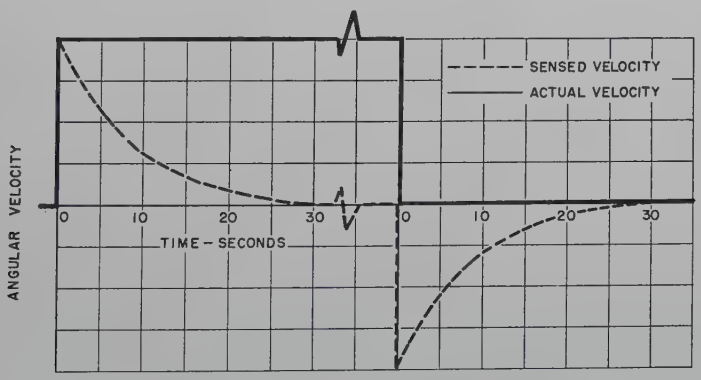

FIGURE 2. Sensed angular velocity for step velocity input

This assumed velocity input is represented by the solid lines in the diagram. The dotted lines represent the displacement of the endolymph and the velocity signal corresponding to this displacement. The effect of the first term of the second member of the equation can be neglected at the scale of the diagram.

Calculations indicate that this term has the effect of delaying the abrupt rise of the fluid displacement by .015 sec.

The return of the endolymph and of the cupula to normal position following abrupt displacement, neglecting as mentioned above the acceleration term of the equation, is given by

$$\theta = \theta_0 e^{-t/\tau}, \tag{4}$$

where

θ = displacement of the endolymph from neutral position at a time t following abrupt acceleration or deceleration,

θ_0 = initial displacement of the endolymph at the end of acceleration period,

$\tau = \frac{c}{K}$ = the time constant of the semicircular canals,

t = time from the end of abrupt acceleration or deceleration,

e = base of natural logarithm.

The curves show that if the body is accelerated abruptly to a constant velocity, the canals will give, immediately following the acceleration, a signal normally corresponding to this velocity. The signal then decays to zero according to equation (4) even though the body is maintained at a constant velocity. If rotation is now suddenly stopped the same phenomenon takes place, except that the endolymph is displaced in the opposite direction. During the return of the fluid and of the cupula to normal position the canals issue velocity signals even though the head is stationary.

These false signals give rise to conflicting information and the well-known post rotational phenomena. They explain the sensation reported in the last sentence of the quotation from Clark and Graybiel. They do not explain, however, the illusory displacement of the sound source. A hypothesis of the cause of displacement must be formulated.

The Role of the Vestibule in Preset Movements. It is known that certain movements of the body can be carried out without the continuous assistance of external senses. This type of control has been discussed, and it is shown (Mayne, 1951) that such movements are carried out in closed loop, with preset input in the neural centers. Such movements are said to be of the preset or autopilot type, to distinguish them from those movements which may be carried out under continuous control of external senses. In the case of a preset movement of the hand, the input in the neural centers may be matched with signals of the proprioceptive senses for closed loop operation. In gross body movements the neural input must be matched

with signals from an absolute spatial reference. It is believed that the vestibule supplies such a reference by a process of double integration (Mayne, 1950, 1951). The first integration is conceived as taking place in the vestibule proper and the second in the neural centers.

The stability of the movement demands that it be controlled as a function of both the error and the velocity of the movement. The error, as in all servo–systems, is the difference between actual and desired position. It is possible to conceive of a system of spatial representation and of a related operation of the vestibule, so that the above requirements for positional and velocity signals will be consistent with the known behavior of the semicircular canals, with post rotational effects, and the presently discussed audiogyral illusion.

Spatial Representation. Space, and the arrangement of objects about us, is probably represented in the neural centers by means of a digital bank of on-off elements. For the sake of simplicity let us consider, rather, an analog type of representation. Let us suppose that the various points of reference about us are merely spotted on a plane, and let us limit our considerations to a two-dimensional representation. Let us draw on our plane, first, two perpendicular axes, representing the front and back, and right and left axes of the body. The eyes may be assumed to be fixed with respect to this system of axes, and the problem may be further simplified by assuming that the head is rigid with respect to the body.

Now, the necessary information is supplied through the complex process of vision to spot various points of reference on the plane. The representation is easiest if the spatial arrangement has symmetry about two axes, and if these axes coincide with the body axes as represented on the plane. This may be why arrangements constructed by men are usually symmetrical about two axes, and why observers naturally orient themselves so that their body axes coincide with those of the arrangement.

The spatial representation may include any number of reference points or lines. At any instant we can close our eyes and have sufficient stored information to perform any number of tasks: touch the corner of our desk, reach for the telephone, stand up and walk to the door of our office without bumping into a chair, etc. For the purpose of this discussion we can reduce the information spotted on the reference plane to a set of two perpendicular axes, one set corresponding to the points of the compass, or to the main axes of symmetry of a room where we may be located, and another set corresponding to the body axes. Our reference plane will then have two sets of axes: an absolute set referred to space and a relative set referred to the body.

If at any time we turn our body with respect to our surroundings, the relative axes must be rotated with it, but the absolute system must remain fixed in space. If, for instance, we sit on our revolving desk chair with our eyes closed and someone turns us 90 degrees, we still can touch our nose and therefore must be aware that it has changed its location in space, or that it is in the same position with respect to our body. At the same time we can still reach for the telephone on our desk, so that we must be aware that it has remained fixed in space, or has changed its location with respect to our body.

Proper orientation requires, therefore, that the absolute axes be rotated with respect to the plane of representation and to the relative axes during rotation of the body. This rotation is believed to be effected as a function of signals from the semicircular canals. It is evident that there are no planes in the neural centers where space may be represented as sketched out above. However, an equivalent digital system, where the coordinates of a point would be represented by association with two numbers, can be easily conceived.

A desired rotation of the body can then be preset by calling for a certain angle between relative and absolute axes as an input. The movement may then be controlled by the difference between this input and the actual angles between the two sets of axes. This difference would then correspond to the positional control function. A value proportional to rate should be added to this positional function for stable and rapid control. Such a rate could be obtained by differentiating the positional information. It appears to be obtained, however, by a more direct method. The relative system of axes seems to be shifted ahead by an amount proportional to the velocity indicated by the semicircular canals, so that the error signal would then contain both position and rate terms.

Computation of the Time Constant of the Semicircular Canals. If the foregoing hypothesis is correct, the relative axes of reference of a subject are shifted in direct proportion to the velocity indicated by the semicircular canals and in the direction of this velocity. A sound referred to this system of axes would then appear to be shifted in a direction opposite to that of rotation. The shift is not a function of actual velocity, but of the signal issued by the semicircular canals, as discussed earlier. In the case of the post rotational period, it has been shown that the signal is given by the exponential expression (4), and it therefore should be expected that the sound shift would follow a similar expression. Further, a plot of the shift vs. time should make it possible to compute the time constant of the semicircular canals.

For the purpose of analysis, the data from the six runs in the study by Clark and Graybiel were averaged. There are differences between individual runs which may be significant, and the tests should be repeated with more elaborate equipment. But, since each run represents averages of a large number of individual tests, there appears to be little possibility of interpreting these differences. A closer attention to individual test variance in future experimentation may provide further clues to some of the details of the system of spatial reference.

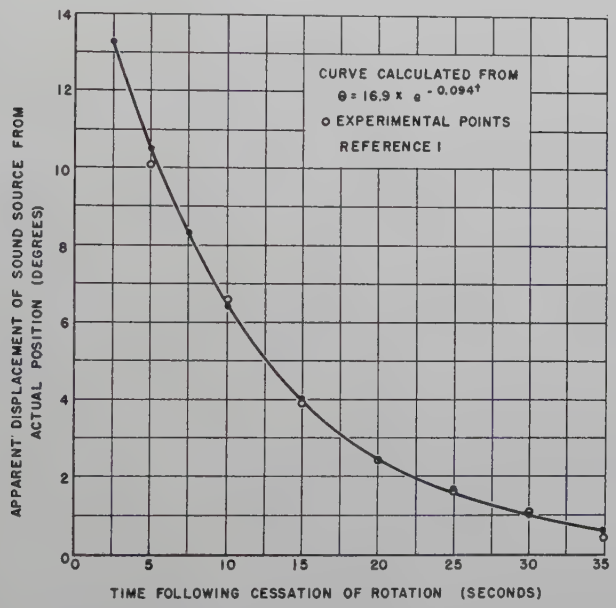

FIGURE 3. Experimental and calculated shift of sound during post-rotational period

As the case may be, the curve of the average disorientation with respect to time is approximated very closely by an exponential where $1/\tau = 0.094$ 1/sec.

Figure 3 shows the agreement between the experimental and the fitted curve. The value of $1/\tau = 0.094$ computed here for human canals agrees reasonably well with that of 0.120 computed in the previous paper, using results (Mowrer, 1935) from experiments on the nystagmic latency of pigeons.

It may be pointed out that objects observed visually do not experience the same illusionary shift of position that occurs in the case of sounds. This may be explained by the nystagmus of the eyes, which compensate for the shift of axes, and in fact overcompensate for such a shift, as shown

by other experimental data. This overcompensation is probably for the purpose of providing for a reaction-time delay in spotting reference points in the neural centers while the head is in motion. It is hoped to present a discussion of nystagmus in a later paper.

Conclusion. The above results lend weight to the following views regarding the semicircular canals and the system of spatial reference.

1. The semicircular canals are a structure defined by equation (2).
2. The values of the time constant of the semicircular canals for man and pigeons are similar.
3. The time constant of the semicircular canal has an approximate value between 0.094 and 0.120 1/sec.
4. The semicircular canals give a signal proportional to the angular velocity of the head within certain limits of velocities and frequencies.
5. The semicircular canals give rate information to stabilize certain body movements.
6. In preset movements rate information combined with positional information is obtained by comparison of an absolute system with a relative system of axes in the neural centers. The absolute system of axes is maintained fixed in space while the relative system is shifted ahead in the direction of the rotation, as a function of information supplied by the semicircular canals.
7. This mechanism explains the audiogyral illusion.

LITERATURE

Clark, Brant and Ashton Graybiel. 1949. "The Effect of Angular Acceleration on Sound Localization: The Audiogyral Illusion." *Jour. Psychol.*, **28,** 235–44.

Mayne, Robert. 1950. "The Dynamic Characteristics of the Semicircular Canals." *Jour. Comp. and Physiol. Psychol.*, **43,** No. 4, 309–19.

———. 1951. "Some Engineering Aspects of the Mechanism of Body Control." *Electrical Engineering,* **70,** No. 3, 207–12.

Mowrer, O. H. 1935. "Nystagmic Response of Pigeons to Constant Acceleration." *Jour. Comp. and Physiol. Psychol.*, **19,** 177–93.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

"IGNITION" PHENOMENA IN RANDOM NETS

ANATOL RAPOPORT
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

The spread of excitation in a "random net" is investigated. It is shown that if the thresholds of individual neurons in the net are equal to unity, a positive steady state of excitation will be reached equal to γ, which previously had been computed as the weak connectivity of the net. If, however, the individual thresholds are greater than unity, either no positive steady state exists, or two such states depending on the magnitude of the axone density. In the latter case the smaller of the two steady states is unstable and hence resembles an "ignition point" of the net. If the initial stimulation (assumed instantaneous) exceeds the "ignition point," the excitation of the net eventually assumes the greater steady state.

Possible connections between this model and the phenomenon of the "preset" response are discussed.

Consider a random net $\mathfrak{N}$ such as was described by R. Solomonoff and A. Rapoport (1951), i.e., an aggregate of neurons, each emitting a axones which synapse randomly on other neurons in the aggregate. Let the threshold of each neuron be 1. Furthermore, let the postulate of "quantized time" hold, that is, firings occur only at multiples of a unit of time, which is taken to be the synaptic delay, assumed constant throughout the net.

Now let a subset of the neurons of $\mathfrak{N}$ be stimulated at $t = 0$. We wish to follow the "expected" course of excitation in $\mathfrak{N}$.

If $x(t)$ is the number of neurons firing at the instant t (an integer), then, following arguments similar to those developed in the earlier paper (*loc. cit.*), we obtain the equations

$$x(t+1) = N\left[1 - \left(1 - \frac{1}{N}\right)^{ax}\right]; \tag{1}$$

$$\frac{x(t+1)}{N} \sim 1 - \exp\frac{-ax}{N}. \tag{2}$$

If a steady state exists, we will denote the number of neurons firing steadily by $N\gamma = x(\infty)$, and obtain an equation in γ identical to equation (20) of the paper mentioned above, namely,

$$\gamma = 1 - e^{-a\gamma}. \tag{3}$$

This result can also be obtained by writing equation (1) as a differential equation, which holds approximately for the case of quantized time:

$$\frac{dx}{dt} = N\left(1 - \exp\frac{-ax}{N}\right) - x(t). \tag{4}$$

Here the right side represents the difference between the expected number of neurons firing at $t+1$ and the number firing at t. But that is the rate of change of x per unit time. Hence it is equal to the left side. Separating the variables, we obtain

$$\frac{dx}{N\left(1 - \exp\frac{-ax}{N}\right) - x} = dt, \tag{5}$$

whose solution for t in terms of x is obtained by a quadrature:

$$t = \int_{x(0)}^{x} \frac{d\xi}{N\left(1 - \exp\frac{-a\xi}{N}\right) - \xi}. \tag{6}$$

The steady state is obtained by setting (4) equal to zero, which again leads to

$$\gamma = 1 - e^{-a\gamma},$$

where $\gamma = x/N$ at the steady state. If this steady state is actually reached, then the net $\mathfrak{N}$ reaches that level of activity regardless of initial conditions. The steady state is thus determined by the structure of the net and not by the initial stimulation imposed upon it.

Remark. This model can be somewhat modified by assuming a refractory period $\delta > 1$. Suppose, as an example, $1 < \delta < 2$, i.e., no neuron can fire at two consecutive units of time, but a neuron may fire every two units of time. In that case, equation (5) becomes

$$\frac{dx}{dt} = (N - x)\left(1 - \exp\frac{-ax}{N}\right) - x. \tag{7}$$

Accordingly, the spread of excitation is given by

$$t = \int_{x(0)}^{x} \frac{dx}{(N - x)\left(1 - \exp\frac{-ax}{N}\right) - x}, \tag{8}$$

and the steady state by the roots of

$$(N - x)\left(1 - \exp\frac{-ax}{N}\right) - x = 0\,, \tag{9}$$

or

$$(1 - \gamma)(1 - e^{-a\gamma}) = \gamma\,. \tag{10}$$

Solving equation (3) for a in terms of γ, we obtain

$$a = \frac{1}{\gamma}\log\frac{1}{1-\gamma}; \tag{11}$$

while equation (10) gives

$$a = \frac{1}{\gamma}\log\frac{1-\gamma}{1-2\gamma}. \tag{12}$$

Equation (12) has meaning only for $\gamma < \frac{1}{2}$, which is intuitively evident, since under the conditions of the refractory period above we cannot have more than one half of the neurons firing twice in succession. However, it can be seen that as a grows to a moderately large value (5 or 6), γ approaches its asymptotic value of $\frac{1}{2}$ very closely. The situation is essentially similar to the one without the refractory period, except for the reduced limiting value of γ as a function of a.

The "Ignition" Threshold. In the case just considered where the threshold of each neuron is unity there is no lower limit on the amount of initial stimulation which eventually leads to the steady state activity. Theoretically, even the firing of a single neuron eventually propagates so as to bring the net into an active steady state, provided $a > 1$. This is not surprising in view of the fact that for $a > 1$, the number of neurons firing at successive times propagates approximately exponentially in the beginning of the process. The situation is essentially different if the thresholds are greater than unity.

In the previous case we had for the probability that a neuron receives a stimulus at time $t + 1$ the following expression:

$$1 - \left(1 - \frac{1}{N}\right)^{ax(t)} \sim 1 - \exp\frac{-ax}{N}. \tag{13}$$

This result is obtained by the following reasoning. At time t, ax stimuli are being sent out by the firing neurons. The probability that a given neuron is *missed* by all of them is

$$\left(1 - \frac{1}{N}\right)^{ax(t)}. \tag{14}$$

Therefore the probability that it receives at least one of them is given by equation (13). However, we can also reason in another way. The probabilities of receiving no stimulus, exactly one stimulus, exactly two stimuli,

etc. are given successively by the terms of the Poisson distribution

$$\begin{aligned} p(0) &= \exp \frac{-ax}{N} \\ p(1) &= \frac{ax}{1!N} \exp \frac{-ax}{N} \\ &\cdot \\ &\cdot \\ &\cdot \\ p(k) &= \frac{a^k x^k}{k! N^k} \exp \frac{-ax}{N}. \end{aligned} \tag{15}$$

The sum of all these probabilities is, of course, unity. Therefore the successive probabilities of receiving *at least* 1, 2, etc. stimuli are

$$\begin{aligned} P(1) &= 1 - \exp \frac{-ax}{N} \\ P(2) &= 1 - \exp \frac{-ax}{N} \left(1 + \frac{ax}{N}\right) \\ &\cdot \\ &\cdot \\ &\cdot \\ P(k) &= 1 - \exp \frac{-ax}{N} E_{k-1}\left(\frac{ax}{N}\right), \end{aligned} \tag{16}$$

where

$$E_k(z) = \sum_{j=0}^{k} \frac{z^j}{j!},$$

as previously defined in an earlier paper (Rapoport, 1950). We are thus led to the generalizaton of expression (13) for the case where the threshold is $h > 1$, namely, the probability of receiving at least h stimuli at time $t + 1$, when $x(t)$ neurons are firing at time t:

$$P(h) = 1 - \exp \frac{-ax}{N} E_{h-1}\left(\frac{ax}{N}\right). \tag{17}$$

This leads immediately to a generalization of equation (6) for this case, namely,

$$\frac{dx}{dt} = N\left[1 - \exp \frac{-ax}{N} E_{h-1}\left(\frac{ax}{N}\right)\right] - x, \tag{18}$$

whose formal solution is given by

$$t = \int_{x(0)}^{x} \frac{d\xi}{N\left[1 - \exp \frac{-a\xi}{N} E_{h-1}\left(\frac{a\xi}{N}\right)\right] - \xi}. \tag{19}$$

The steady state is given by the roots of

$$1 - e^{-a\gamma}E_{h-1}(a\gamma) - \gamma = 0 . \tag{20}$$

If $h = 1$, equation (20) solved for γ has exactly one positive root for $1 < a < \infty$, as has been shown (Solomonoff and Rapoport, 1951). It will now be shown that if $h > 1$, the corresponding equation (20) has exactly two positive roots, provided a exceeds a certain value, which depends on h.

We write equation (20) as

$$1 - \gamma = e^{-a\gamma}E_{h-1}(a\gamma) . \tag{21}$$

Holding a and h fixed, plot the left and right sides of (21) against γ (Fig. 1). Clearly the abscissae of the intersections of the two curves will de-

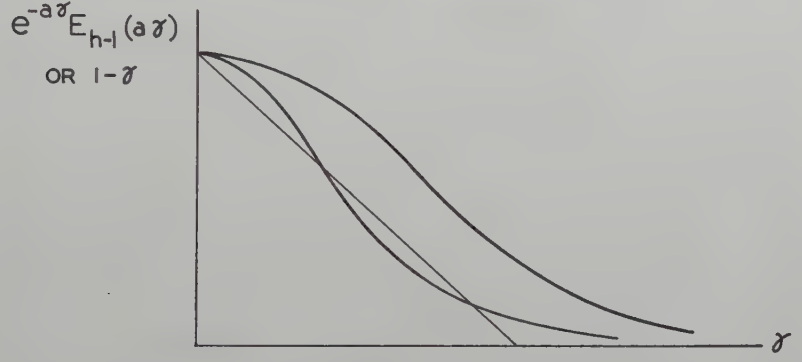

FIGURE 1. The line $1 - \gamma$ and the curves $e^{-a\gamma}E_{h-1}(a\gamma)$ for two values of a plotted against γ. For a sufficiently small value of a (upper curve) no intersection occurs. For sufficiently large values, there are exactly two intersections.

termine the values of γ which will satisfy (21) for a given pair of values of a and h. We note that if $h > 1$, then

$$\frac{d}{d\gamma} e^{-a\gamma}E_{h-1}(a\gamma) = -a e^{-a\gamma}E_{h-1}(a\gamma) + a e^{-a\gamma}E_{h-2}(a\gamma) . \tag{22}$$

At $\gamma = 0$,

$$\frac{d}{d\gamma} e^{-a\gamma}E_{h-1}(a\gamma) = -a + a = 0 . \tag{23}$$

Moreover,

$$\frac{d^2}{d\gamma^2} e^{-a\gamma}E_{h-1}(a\gamma) = a^2 e^{-a\gamma}E_{h-1}(a\gamma) - 2a^2 e^{-a\gamma}E_{h-2}(a\gamma) + a^2 e^{-a\gamma}E_{h-3}(a\gamma) .* \tag{24}$$

* Consistent with the definition of E_h, the value 0 is assigned to E_{-1} (cf. Rapoport, 1950).

If we set the right side of (24) equal to zero, we find that since $a^2e^{-a\gamma} \neq 0$,

$$E_{h-1}(a\gamma) - 2E_{h-2}(a\gamma) + E_{h-3}(a\gamma) = 0\,, \tag{25}$$

which may be rewritten as

$$(E_{h-1} - E_{h-2}) - (E_{h-2} - E_{h-3}) = \frac{a^{h-1}\gamma^{h-1}}{(h-1)!} - \frac{a^{h-2}\gamma^{h-2}}{(h-2)!} = 0\,. \tag{26}$$

The roots of (26) are at $\gamma = 0$ and $\gamma = (h-1)/a$. This means that the graph of $e^{-a\gamma}E_{h-1}(a\gamma)$ has at most two inflection points, namely, at $\gamma = 0$ and at $\gamma = (h-1)/a$. Consider now the slope $e^{-a\gamma}E_{h-1}(a\gamma)$ at $\gamma = (h-1)/a$. Substituting this value into equation (22), we find:

$$\begin{aligned} &\frac{d}{d\gamma}\, e^{-a\gamma}E_{h-1}(a\gamma)\,\Big|_{\gamma=(h-1)/a} \\ &\quad = -a\,e^{-(h-1)}E_{h-1}(h-1) + a\,e^{-(h-1)}E_{h-2}(h-1) \\ &\quad = \frac{-a\,e^{-(h-1)}(h-1)^{h-1}}{(h-1)!} \sim \frac{-a}{\sqrt{2\pi(h-1)}}\,, \end{aligned} \tag{27}$$

where $h > 1$.

Hence $a/\sqrt{2\pi(h-1)}$ is approximately the greatest absolute value of the slope of $e^{-a\gamma}E_{h-1}(a\gamma)$. If this value does not exceed unity, that curve will never cross the line $1 - \gamma$, and equation (20) will have no positive root (cf. Fig. 1). Therefore we have deduced

Lemma 1. If $h > 1$, no positive steady state exists, unless approximately

$$a > \sqrt{2\pi(h-1)}\,. \tag{28}$$

Remark. Inequality (28) is a necessary, but not a sufficient, condition on the axone density of the net for the existence of a positive steady state.

Lemma 2. If $h > 1$, there exist either no positive steady states or exactly two such states.*

Proof. We note that the right side of equation (21) approaches zero monotonically from above. Therefore if that curve has crossed the line $1 - \gamma$ at some point where $\gamma < 1$, it must cross it again at some other point where $\gamma < 1$. The two intersections determine the values of γ, where $dx/dt = 0$, i.e., steady states.

Lemma 3. Let the two positive steady states (if they exist) occur at γ_1 and γ_2 ($\gamma_1 < \gamma_2$). Then $x/N = \gamma_1$ gives an unstable steady state while $x/N = \gamma_2$ gives a stable one.

* We are disregarding the very special case where tangency occurs.

This follows immediately from the inequalities

$$\frac{dx}{dt} < 0, \quad \text{for} \quad \frac{x}{N} < \gamma_1; \quad \frac{dx}{dt} > 0, \quad \text{for} \quad \gamma_1 < \frac{x}{N} < \gamma_2;$$
$$\frac{dx}{dt} < 0, \quad \text{for} \quad \frac{x}{N} > \gamma_2. \tag{29}$$

We thus have the following situation. Suppose a and h are fixed in such a way that two positive steady states exist. Then if the initial stimulation of $\mathfrak{N}$ is such that $x(0) < N\gamma_1$, the activity will die out. If, however, the initial activity exceeds $N\gamma_1$, the activity will increase until steady state $N\gamma_2$ is reached. If the initial activity exceeds $N\gamma_2$, it will *decrease* to that

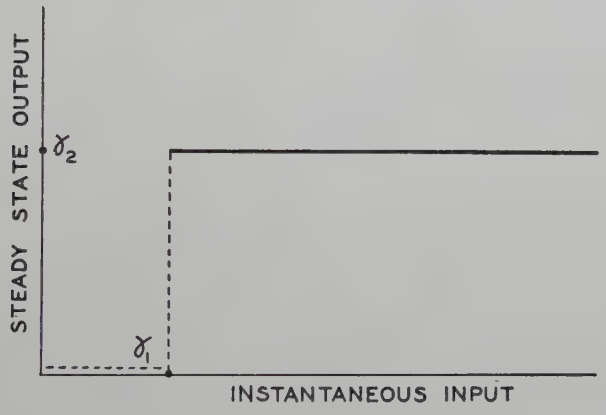

FIGURE 2. The all-or-none character of the steady state of excitation in a random net

steady state. We have here essentially an "ignition" phenomenon, where the threshold of ignition acts as the *overall* threshold of the net. If the initial instantaneous stimulation of the net involves a number of neurons in excess of that threshold (i.e., $N\gamma_1$), the net will "ignite" and will remain active to the extent of the involvement of $N\gamma_2$ neurons in that activity. The input-output curve of such a net is shown in Figure 2. The input, however, must be here understood as an initial instantaneous input and not a continuous one as was the case in our earlier papers dealing with input-output problems. The output, on the other hand, is the steady state output of the net. In terms of these definitions, the net exhibits an ali-or-none behavior.

Relations between the Parameters of the Net. Our treatment has given rise to the following parameters, characterizing the net $\mathfrak{N}$: the axone density a; the individual neuron threshold h, which is an integer designating the number of stimuli which must be "simultaneously" received by a neuron

to elicit a firing; the net threshold γ_1, designating the fraction of the neurons which must be initially stimulated for the net to become permanently active; and γ_2, the fraction of neurons active in the stable positive steady state. We seek relations among these parameters.

It is easy to verify that if $h < h'$, then

$$e^{-a\gamma}E_{h-1}(a\gamma) < e^{-a\gamma}E_{h'-1}(a\gamma) \tag{30}$$

for all values of $\gamma > 0$. Therefore, for fixed a, the curve $e^{-a\gamma}E_{h'-1}(a\gamma)$ will lie wholly above the curve $e^{-a\gamma}E_{h-1}(a\gamma)$, as shown in Figure 3. If

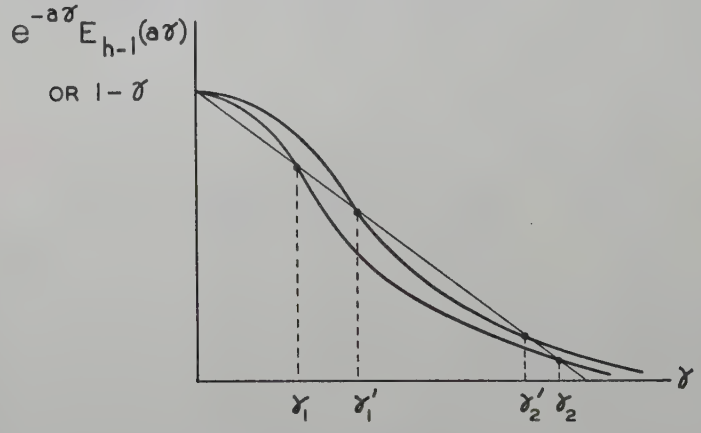

FIGURE 3. The curves $e^{-a\gamma}E_{h-1}(a\gamma)$ for two values of h. The upper curve corresponds to the higher value of h.

γ_1' and γ_2' are the threshold and steady state parameters associated with h', then we shall have

$$\gamma_1' > \gamma_1 ; \qquad \gamma_2' < \gamma_2 . \tag{31}$$

In other words, if the thresholds of the neurons comprising the net $\mathfrak{N}$ are uniformly raised, the overall threshold of the net will be raised and the steady state activity lowered, as, of course, should be the case. Vice versa, a lowering of the individual thresholds will lower the overall net threshold and raise the steady state activity.

We have here a model of what has been called a "preset" response (Mayne, 1951). Suppose $\mathfrak{N}$ is innervated by other centers, $C_1, C_2, \ldots C_k$ in such a way that each neuron of $\mathfrak{N}$ receives an axone from each of the C's. For simplicity of discussion, we still assume that all firings are synchronized and occur once per unit time. Then if the centers $C_1 \ldots . C_k$ are totally activated, and $k < h$, the stimulation of $\mathfrak{N}$ by the C's is subliminal. However, the effect of the continued activation of the C's is to re-

duce the individual thresholds of the neurons of $\mathfrak{N}$ by k units with respect to stimulation from other sources. Thus the amount of response which will occur in $\mathfrak{N}$ as a result of outside stimulation has been "preset" by subliminal stimulation from the centers $C_1 \ldots . C_k$, and this response is again independent (after the steady state has been established) of the magnitude of the outside (superthreshold) stimulus.

We may now generalize this model by supposing that the individual thresholds in $\mathfrak{N}$ are not uniformly distributed. Or we can suppose that the centers C innervate $\mathfrak{N}$ only partially, each center innervating its particular set of neurons in $\mathfrak{N}$. Then a distribution of activity in the centers C amounts to a new distribution of individual thresholds in $\mathfrak{N}$.

Let the fraction f_1 of the neurons in $\mathfrak{N}$ have threshold 1, the fraction f_2 have threshold 2, etc. If some values of the threshold are not represented in the net, the corresponding f's can be taken to be zero. Hence formally the indices of the f's can range from 1 to infinity. We suppose that the neurons of each type are spread at random through the net. Then we shall have the spread of excitation in $\mathfrak{N}$ governed by the following equation

$$\frac{dx}{dt} = N \sum_{j=1}^{\infty} f_j \left[1 - \exp \frac{-ax}{N} E_{j-1}\left(\frac{ax}{N}\right)\right] - x . \tag{32}$$

The steady state will be given by the solutions of

$$\sum_{j=1}^{\infty} f_j \left[1 - e^{-a\gamma} E_{j-1}(a\gamma) \right] - \gamma = 0 .$$

However, the following identities hold:

$$\begin{aligned}
1 - e^{-a\gamma} E_0(a\gamma) &= 1 - e^{-a\gamma} \\
1 - e^{-a\gamma} E_1(a\gamma) &= 1 - e^{-a\gamma} - \frac{a\gamma}{1!}\, e^{-a\gamma} \\
&\cdot \\
&\cdot \\
&\cdot \\
1 - e^{-a\gamma} E_{h-1}(a\gamma) &= 1 - e^{-a\gamma} - \frac{a\gamma}{1!}\, e^{-a\gamma} - \ldots . \frac{a^{h-1}\gamma^{h-1}}{(h-1)!}\, e^{-a\gamma} .
\end{aligned} \tag{33}$$

Moreover, if h is the highest threshold in $\mathfrak{N}$, we have

$$\sum_{j=1}^{h} f_j = 1 . \tag{34}$$

Multiplying the right sides of (33) by appropriate factors f_j and adding,

we obtain

$$\frac{dx}{dt} = 1 - e^{-a\gamma} - F_1 \frac{a\gamma}{1!} e^{-a\gamma} - F_2 \frac{a^2\gamma^2}{2!} e^{-a\gamma} - \ldots . F_{h-1} \frac{a^{h-1}\gamma^{h-1}}{(h-1)!} e^{-a\gamma}, \quad (35)$$

where

$$F_i = \sum_{j=i+1}^{h} f_j = 1 - \sum_{j=1}^{i} f_j . \quad (36)$$

The steady states are obtained as before by setting $dx/dt = 0$.

We shall investigate the special case where two classes of neurons are involved with thresholds $h = 1$ and $h = 2$ occurring with frequencies f and $1 - f$ respectively. This situation may be interpreted as involving a net with all neuron thresholds equal to 2 but with a fraction f of the neurons *subliminally* stimulated. The steady state equation is

$$1 - e^{-a\gamma} - (1 - f)\, a\gamma e^{-a\gamma} - \gamma = 0 . \quad (37)$$

We have seen that for $f = 1$, $a > 1$, equation (37) determines exactly one positive stable steady state. For $f = 0$, a must exceed a certain value for a positive stable state to exist. Suppose the condition is satisfied. Then the threshold γ_1 and the steady state fraction γ_2 are continuous functions of f, given implicitly by equation (37). By varying f continuously, the threshold γ_1 and the steady state γ_2 may be made to vary continuously also. Physiologically, this may mean that the activity of a center C, whose axones innervate the neurons of $\mathfrak{N}$ subliminally, determines the overall threshold ("ignition point") of $\mathfrak{N}$ and the steady state excitation which $\mathfrak{N}$ will attain if stimulated sufficiently by an initial outside stimulus.

Such a mechanism may possibly underlie a "preset" response, where an "estimate" of the desired "intensity" of some act seems to be made in *advance*, so that the act can be initiated by a trigger stimulus and then be performed automatically with the desired intensity. Possible examples are estimating the distance to be jumped, the pitch of a tone to be sung, etc. For experimental results with a possible bearing on preset responses, see the references given in the paper mentioned above of R. Mayne.

This investigation is part of the work done under Contract No. AF 19(122)-161 between the U.S. Air Force Cambridge Research Laboratories and the University of Chicago.

LITERATURE

Mayne, Robert. 1951. "Some Engineering Aspects of the Mechanism of Body Control." *Electrical Engineering*, March.

Rapoport, Anatol. 1950. "Contribution to the Probabilistic Theory of Neural Nets: II. Facilitation and Threshold Phenomena." *Bull. Math. Biophysics*, **12**, 187–97.

Solomonoff, R. and A. Rapoport. 1951. "Connectivity of Random Nets." *Bull. Math. Biophysics*, **13**, 107–17.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

DETERMINATION OF DIFFUSION AND PERMEABILITY COEFFICIENTS IN MUSCLE

I. Opatowski and George W. Schmidt
Committee on Mathematical Biology
The University of Chicago

A theory is presented for the study of diffusion in heterogeneous tissue-like structures. It is applicable to a common type of measurement in which the change of the amount of substance remaining in the tissue is determined as the substance diffuses from the tissue into an adjacent medium, for instance, Ringer's solution. The main objective of this paper is to obtain a method for the calculation of the diffusion coefficient in the intercellular space and of the permeability coefficients between this space and the cells, based on the type of measurement mentioned above. Although the fundamental ideas upon which the theory is based are applicable to any type of tissue, the formulae derived are limited to the case in which the cells form a flat bundle of parallel fibers. The theory is applied to the experimental results of E. J. Harris and G. P. Burn on diffusion of sodium in the sartorius muscle of the frog. We find that if we know the ratio of the cellular and intercellular volumes of the muscle the ratio of the equilibrium concentrations of sodium outside and inside the cells can be determined. A very simple mathematical analysis of the experimental relation between the amount of substance diffusing out of the muscle and the time of diffusion gives us this ratio. The ratio of the equilibrium sodium concentrations in the case of the sartorius frog muscle is between about 10 and 30, depending on the muscle used. The same mathematical analysis makes it possible to obtain the permeability coefficients of muscle fibers through simple calculations, if their sizes are known. The permeability coefficients for the experimental work mentioned above using sodium are 1.25 to 11.5×10^{-8} cm/sec for the flow into the fibers and 3.2 to 16×10^{-7} cm/sec for the flow in the opposite direction. The determination of the diffusion coefficient in the intercellular space is more laborious and yields only an order of magnitude: 10^{-6} cm^2/sec.

1. *The objective of the paper and its fundamental equations.* Harris and Burn have published an experimental and theoretical study on diffusion in tissue in which an attempt is made to characterize the properties of a tissue by considering the process of diffusion in the intercellular space separately from the permeability through the cell membrane. The experimental technique used by these authors (1949) consisted of following the time dependence of the total amount of a substance existing in a tissue and diffusing outside of it into a solution of known concentration. The

diffusion coefficient in the intercellular space and the permeability coefficients between the latter and the cells were obtained from a mathematical analysis of that time dependence. This seems to be the first time that instead of a determination of an average diffusion or permeability coefficient of the tissue as a whole a more accurate analysis reflecting its heterogeneous structure has been attempted.

It is the purpose of the present work to develop a theory based on the fundamental laws of diffusion and permeability. We shall also take into account the heterogeneous structure of the tissue and show the applicability of the theory to the type of experimental result obtained by Harris and Burn.

Consider an idealized geometry of a tissue slice (Fig. 1) of unit thickness in the direction perpendicular to the diagram. Let the length of the slice in the plane of the diagram be infinite, let the cells all be of equal size and

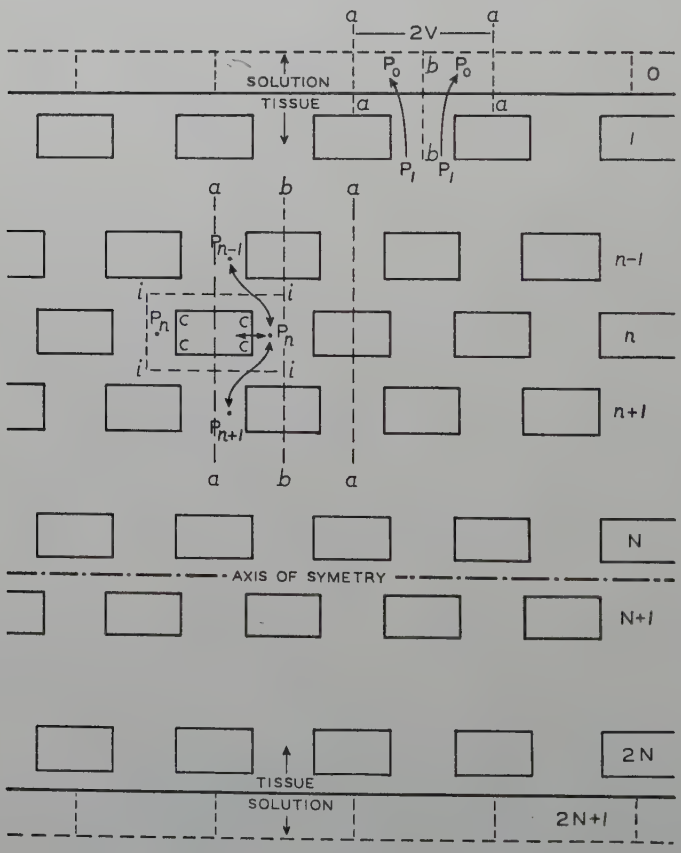

FIGURE 1. An idealized geometry of a tissue slice upon which the theory is based

shape and have a fiber-like prismatic structure of cross-section cccc with axes parallel to each other arranged in a regular fashion in parallel rows 1, 2, ... , n, ... , $2N$. Consider the case in which the diffusing substance is initially distributed in a homogeneous fashion throughout the whole intercellular volume. Let the concentration of the substance be the same initially in all cells. At $t = 0$ the tissue is immersed in a solution and the substance diffuses into the surrounding medium through the two end layers $n = 1$ and $n = 2N$. We assume that the total number of layers of the tissue is even. Then the study can be limited to one-half of the tissue, since an axis between the layers N and $N + 1$ is an axis of symmetry (Fig. 1) through which no diffusion occurs. Consequently the diffusion in one-half of the tissue (Fig. 1), i.e., for $n = 1$ to $n = N$, is the same as would occur if this part of the tissue had its layer $n = 1$ in contact with the solution and the layer $n = N$ in contact with a solid into which there could be no diffusion. Thus we see that the determination of diffusion in one-half of the layer shown in Figure 1 actually solves two different problems, both corresponding to the previously mentioned measurements of Harris and Burn. The theory developed here could also be extended to the case in which the total number of layers of the tissue, as represented in Figure 1, is odd, if we consider that the substance flows from the middle layer into both adjacent layers at an equal rate.

The idealization of the tissue as represented in Figure 1 makes the structure a periodic one, the strip between the axes aa representing a full period and the axis bb being an axis of symmetry within the period. Naturally the "cells" of Figure 1 are rectangular only for convenience of discussion; mathematical considerations do not imply any particular shape for the cell. The axes aa are also axes of symmetry; consequently, no substance is diffusing across the lines aa and bb. The diffusion goes on in each strip aabb independently of the other strips. In general, the diffusing substance is distributed throughout the whole tissue. We assume that insofar as the overall diffusion process is concerned the distribution of the substance in the intercellular spaces may be approximated with sufficient accuracy by using the average distribution over the intercellular space associated with any cell instead of the actual distribution. This assumption can be formulated in the following manner.

Consider a cell cccc of the nth row. Let C_n be the average concentration in one-half of this cell, which is determined by the boundary of the latter and by the axis aa. Because of symmetry C_n is also the average concentration in the whole cell. We assume that the behavior of the cell in the tissue is described with sufficient accuracy by the magnitude of this average con-

centration alone. We consider each cell to be surrounded by an adjacent intercellular space with its external boundary at iiii (Fig. 1). Let I_n be the average concentration in such a space of the nth row. The average concentration in each half of the intercellular space bounded by aa is also I_n for reasons of symmetry. Whenever concentrations have to be assigned to well defined points, we will imagine the concentrations I_n exist at certain points P_n inside the intercellular spaces. For reasons of symmetry there are two such points around each cell. On the basis of this scheme diffusion flow connects each half of a cell directly only with the adjacent half of the intercellular space. Consequently the change of the amount of substance in the cell is, per unit of time,

$$\frac{A\,dC_n}{dt} = S\,(k_I I_n - k_c C_n)\,, \qquad n = 1, 2, \ldots, N\,, \tag{1}$$

where A is the cross-sectional area of the cell, i.e., the volume of the cell, since the tissue is assumed to be of unit length in the direction perpendicular to the diagram; S is the area of the cell membrane, k_I and k_c its permeability coefficients. The subscript n is limited to $1, \ldots, N$ because we consider only one-half of the tissue up to the axis of symmetry. We shall now consider that part of an intercellular space of the nth row which lies between the axes aa and bb. Three diffusion flows exist between this space and its surroundings as indicated by arrows in Figure 1. Two of these flows are from or into the adjacent intercellular spaces of the rows $n - 1$ and $n + 1$, one is from or into the surrounded cell. The average concentration gradient in the diffusion from the intercellular space surrounding P_{n-1} to the intercellular space surrounding P_n can be taken with good approximation to be

$$\frac{I_{n-1} - I_n}{L}, \qquad (n = 2, \ldots, N - 1)\,,$$

where L is a length which does not differ by very much from the length of an arc like $P_{n-1}P_n$ (Fig. 1). This is about a linear dimension of the cell plus the intercellular distance. We assume that L is the same for any two adjacent rows, that is, L is independent of n. The number of molecules which diffuse per unit time from the part of the intercellular space surrounding P_{n-1} to the part surrounding P_n can be written in the form:

$$DW\,\frac{I_{n-1} - I_n}{L}, \qquad (n = 2, \ldots, N)\,,$$

where W is an average width of this diffusion flow and D is the diffusion coefficient in the intercellular space. The length $2W$ does not differ greatly

from the average distance of two neighboring cells. Using these symbols the change of the number of molecules in the intercellular space bounded by iiii and cccc is, per unit of time,

$$\frac{B d I_n}{dt} = \frac{2DW}{L}(I_{n-1} - 2I_n + I_{n+1}) - S(k_I I_n - k_c C_n) \quad (n = 2, \ldots, N-1), \quad (2)$$

where B is the cross-sectional area of that intercellular space, i.e., its volume, because we consider a slice of tissue of unit thickness. In numerical calculations B is often taken to be somewhat smaller than that volume in order to take into account the solid substances that it contains into which practically no diffusion flow occurs. Since no substance is flowing across the axis of symmetry we have for $n = N$, instead of equation (2),

$$B d I_N dt = \frac{2DW}{L}(I_{N-1} - I_N) - S(k_I I_N - k_c C_N). \quad (3)$$

We assume that the concentration of the diffusing substance is negligible in the external solution. Then the concentration gradient between layer 1 and the boundary of the tissue is $I_1/(L/2)$ and, instead of equation (2), we obtain for $n = 1$:

$$\frac{B d I_1}{dt} = \frac{2DW}{L}(I_2 - 3I_1) - S(k_I I_1 - k_c C_1). \quad (4)$$

The diffusion problem outlined corresponds to the following initial conditions for $I_n(t)$ and $C_n(t)$:

$$I_n(0) = I, \left(\frac{dC_n}{dt}\right)_{t=0} = 0, \text{ i.e., } k_c C_n(0) = k_I I_n(0), \text{ for } n = 1, \ldots, N. \quad (5)$$

These conditions imply an initially constant concentration I in the intercellular spaces and a constant concentration in the cells in equilibrium with those spaces. The cellular concentration C_n in any of the layers can be calculated as soon as the intercellular concentration I_n in the same layer is known. In fact, from equation (1), we have for $n = 1, \ldots, N$:

$$\epsilon(t) C_n(t) = \frac{k_I}{k_c} I + k_I \frac{S}{A} \int_0^t \epsilon(t) I_n(t) dt, \quad (6)$$

where $\epsilon(t) = \exp(k_c St/A)$.

Take a portion of the tissue of *unit thickness* consisting of N layers ($n = 1, \ldots, N$ or $n = n + 1, \ldots, 2N$) with M cells and M intercellular spaces in each layer. For this portion of the tissue let

$a = MNA$ = total cross-sectional area of the cells = total cellular

volume,

$\beta = MNB$ = total cross-sectional area of the intercellular spaces = total intercellular volume involved in the diffusion process,

$\sigma = MNS$ = total surface of the cell membranes,

$\lambda = NL$ = average length of the diffusion path within the intercellular spaces across the whole portion of the tissue from one boundary to the axis of symmetry. The quantity 2λ is likely to be somewhat larger than the size of the whole tissue in the direction aa.

$X_n = \alpha C_n/N$ = amount of substance in all the cells of the nth row,

$Y_n = \beta I_n/N$ = amount of substance in all the intercellular spaces of the nth row,

$Q = \sum_{n=1}^{n=N} (X_n + Y_n)$ = total amount of the substance in the portion of tissue considered.

Using these symbols equations (1) and (2) can be transformed into:

$$\begin{aligned} \frac{dX_n}{dt} &= \varphi_2 Y_n - \varphi_1 X_n, \qquad (n = 1, \ldots, N), \\ \frac{dY_n}{dt} &= \xi (Y_{n-1} - 2Y_n + Y_{n+1}) - \varphi_2 Y + \varphi_1 X_n, (n = 2, \ldots, N-1), \end{aligned} \tag{7}$$

where $\varphi_1 = k_c\sigma/\alpha$, $\varphi_2 = k_I\sigma/\beta$, $\varphi = 2DWM/(\beta\lambda)$, $\xi = \varphi N^2$, $X_n(0) = (\varphi_2/\varphi_1)Y/N$, $Y_n(0) = Y/N$, $Y = \beta I$ is the total amount of substance existing initially in all the intercellular spaces, and $(\varphi_2/\varphi_1)Y$ is the total amount of substance existing initially in all the cells of the portion of tissue $n = 1, \ldots, N$ or $n = N+1, \ldots, 2N$. Equations (3) and (4) yield two other relations obtainable from equation (7) if we put $Y_0 = -Y_1$, $Y_{N+1} = Y_N$ by definition.

2. *Solution of the equations; total amount of the substance in the tissue.* To solve the system of equations derived at the end of the previous section we shall apply the Laplace transformation, that is, associate with each function $F(t)$ its Laplace transform $f(s)$ as defined by

$$f(s) = \int_0^\infty e^{-st} F(t)\, dt.$$

We will use the following type of symbolism:

$$f(s) = \mathfrak{L}\{F(t)\}, \qquad F(t) = \mathfrak{L}^{-1}\{f(s)\}.$$

If we put $x_n(s) = \mathfrak{L}\{X_n(t)\}$ and $y_n(s) = \mathfrak{L}\{Y_n(t)\}$ we obtain:

$$\begin{aligned} s x_n - X_n(0) &= \varphi_2 y_n - \varphi_1 x_n \qquad (n = 1, 2, \ldots, N) \\ s y_n - Y_n(0) &= \xi (y_{n-1} - 2y_n + y_{n+1}) \\ &\quad - \varphi_2 y_n + \varphi_1 x_n \qquad (n = 2, 3, \ldots, N-1). \end{aligned} \tag{8}$$

The last equation is valid for $n = 1$ and $n = N$ if we define $y_0 = -y_1$ and $y_{N+1} = y_N$. Taking into account the expressions of $X_n(0)$ and $Y_n(0)$ we obtain for $n = 1, 2, \ldots, N$, using the above definitions of y_0 and y_{N+1},

$$\begin{aligned} X_n = \frac{\varphi_2}{\varphi_1} \frac{\varphi_1 y_n + N^{-1} Y}{s + \varphi_1} [1 + \varphi_2 (s + \varphi_1)^{-1}] (s y_n - N^{-1} Y) \\ = \xi (y_{n-1} - 2y_n + y_{n+1}). \end{aligned} \tag{9}$$

The solution of equation (9) for $n = 1, 2, \ldots, N$ is:

$$y_n = \frac{Y}{SN}\left[1 - 2\,\frac{z^n + z^{2N+1-n}}{(1+z)(1+z^{2N})}\right], \tag{10}$$

where $z = u + \sqrt{u^2 - 1}$; or

$$u = \frac{1}{2}\left(z + \frac{1}{z}\right), \qquad u = 1 + \frac{s}{2\xi}\left(1 + \frac{\varphi_2}{s + \varphi_1}\right). \tag{11}$$

The quantities u and y_n do not change if z is changed in them into $1/z$, because the coefficients of y_{n-1} and y_{n+1} in the difference equation for y_n are the same.

Let $Q(t)$ be the amount of substance existing at any time t in the portion of tissue from $n = 1$ to $n = N$. The Laplace transform of $Q(t)$ is

$$\begin{aligned} q(s) &= \mathfrak{L}\{Q(t)\} = \sum_{n=1}^{n=N} (x_n + y_n) = \frac{Y}{s}\left(1 + \frac{\varphi_2}{\varphi_1}\right)[1 - h(s)], \\ h(s) &= \left(1 + \frac{\varphi_2}{\varphi_1}\right)^{-1} [1 + \varphi_2 (s + \varphi_1)^{-1}] T, \\ T &= \frac{2}{N} \frac{z(z^{2N} - 1)}{(z^2 - 1)(z^{2N} + 1)}. \end{aligned} \tag{12}$$

The total amount of substance initially existing in the portion of tissue considered is

$$Q(0) = Y\left(1 + \frac{\varphi_2}{\varphi_1}\right).$$

Consequently $G(t) = Q(t)/Q(0)$ is *the fraction of the initial amount of the substance which at the moment* t *still exists in the tissue.* We obtain for the

Laplace transform of $G(t)$:

$$g(s) = \mathfrak{L}\{G(t)\} = \frac{1-h(s)}{s}.$$

An easy check of the calculations made up to this moment is possible. The function u tends to 1 as $s \to 0$; consequently z tends also to 1. Therefore, $h(0) = 1$ and, we find, by an application of a Tauberian theorem of the Laplace transformation (Doetsch, 1937, p. 208, theorem 2):

$$\mathfrak{L}^{-1}\left\{\frac{h(s)}{s}\right\} \to 1 \qquad \text{when} \qquad t \to \infty .$$

Consequently $G(\infty) = 0$, as it should because at $t = \infty$ no substance is left in the tissue. When $s \to \infty$, $u \to \infty$ and $z \to \infty$. Consequently $h(s) \to 0$ and $g(s)$ behaves at infinity as $1/s$. Therefore $G(0) = 1$, as it should by the meaning of the function $G(t)$ itself.

We now calculate $\mathfrak{L}^{-1}\{T(s)\}$. The poles of T considered as a function of z are the zeros of the denominator of T that are not zeros of its numerator. Consequently, these poles are roots of the equation

$$z^{2N} + 1 = 0 , \tag{13}$$

i.e., they are the following values of $z = z_n$:

$$z_n = \exp(i\theta_n) \quad \text{where} \quad \theta_n = \frac{(2n-1)\pi}{2N} \quad \text{with} \quad n = 1, 2, \ldots, N . \tag{14}$$

In order to discuss the singularities of T considered as a function of u we need the following theorem: If a function $F(z)$ has a simple pole at $z = z_n$, and if $u = u_n$ is a simple root of the equation $z(u) = z_n$, where $z(u)$ is an analytic function of u at $u = u_n$, then $u = u_n$ is also a simple pole of $F[z(u)]$ considered as a function of u. The proof of this theorem is easy. In fact,

$$F(z) = \frac{K}{z - z_n} + \Phi(z) ,$$

$$z - z_n = (u - u_n)\Psi(u) ,$$

where $\Phi(z)$ and $\Psi(u)$ are analytic functions at $z = z_n$ and $u = u_n$ respectively, and K and $\Psi(u_n)$ are two finite constants different from zero. Therefore

$$(u - u_n) F[z(u)] = \frac{K}{\Psi(u)} + (u - u_n)\Phi[z(u)] .$$

Since $\Phi[z(u_n)]$ is finite, $\Psi(u_n)$ and K are both not equal either to zero or infinity, we see from the last equation that $u = u_n$ is a simple pole of $F[z(u)]$ with the residue $K/\Psi(u_n)$. This proves the theorem.

To each value of $z = z_n$ corresponds by the first equation of (11) one,

and only one, value of u which is $u_n = \cos\theta_n$. The function $z(u)$ is analytic at all $u = u_n$. Consequently $u = u_n$ are *simple poles of* T *considered as a function of* u. Since

$$\cos\theta_n = \cos\theta_{2N-n+1}\,, \tag{15}$$

the poles of T considered as a function of z [see equation (14)] can be grouped in pairs so that each pair gives rise to a single pole of T considered as a function of u. Consequently, in dealing with this latter function *we can limit the range of the subscript* n *to* n = 1, . . . , N. It may be easily seen that beyond $u = u_n$ there are no other singularities of T since these could arise only out of the singularities of the function $z(u)$ which are $u = \pm\infty$, ± 1 and correspond to $z = 0, \pm\infty, \pm 1$. A direct inspection shows, however, that these are not singularities of T.

We proceed now to discuss T as a function of s. The relation between s and u is given by the second equation of (11) which can also be written in the form:

$$s^2 + (\varphi_1 + \varphi_2 + \nu\xi)\, s + \nu\varphi_1\xi = 0\,, \tag{16}$$

where $\nu = 2 - 2u$. Two values of s which we call s_n correspond to each value of $u = u_n$. They are the roots of the following equation in s_n:

$$s_n^2 + (\varphi_1 + \varphi_2 + \nu_n\xi)\, s_n + \nu_n\varphi_1\xi = 0\,, \tag{17}$$

where $\nu_n = 2 - 2\cos\theta_n$ and $n = 1, \ldots, N$. The roots of equation (17) are always simple, both real and negative. The larger of these roots will be indicated by s_n^+, the smaller by s_n^-. It is seen from the second equation of (11) that the function $u(s)$ is analytic at all $s = s_n$, and, therefore, $s = s_n^+$ and $s = s_n^-$ are *simple poles of* T *considered as a function of* s. It may be easily seen that beyond $s = s_n^+, s_n^-$ there are no other singularities of T, since these could arise only out of the singularities of the function $u(s)$ which are $s = -\varphi_1, \pm\infty$ and correspond to $u = \pm\infty$. But the latter are not singularities of T.

We can now apply the Heaviside partial fraction expansion as follows:

$$\mathfrak{L}^{-1}\{T\} = \sum_{s_n = s_n^+, s_n^-} \sum_{n=1}^{n=N} T_n\,,$$

$$T_n = \frac{2}{N}\,\frac{z_n(z_n^{2N} - 1)}{z_n^2 - 1}\,\exp(s_n t)\lim_{s\to s_n}\frac{s - s_n}{z^{2N} + 1}\,.$$

Since $z_n^{2N} = -1$ and, consequently, $z_n^{2N-1} = -z_n^{-1}$, we obtain:

$$T_n = \frac{2}{N^2}\,\frac{z_n}{z_n - z_n^{-1}}\,\exp(s_n t)\left(\frac{ds}{dz}\right)_{s=s_n}\,.$$

If we put $p = \varphi_1 + s$, equation (5) becomes:

$$p^2 - (\varphi_1 - \varphi_2 - \nu\xi)\, p - \varphi_1\varphi_2 = 0 .$$

From this we find:

$$\frac{ds}{dz} = \frac{dp}{d\nu}\frac{d\nu}{du}\frac{du}{dz} = \frac{\xi}{z^2}\,\frac{z^2-1}{1+\dfrac{\varphi_1\varphi_2}{p^2}} .$$

From the last expression of T_n we obtain:

$$T_n = 2\varphi\left(1+\frac{\varphi_1\varphi_2}{p_n^2}\right)^{-1}\exp(s_n t) . \tag{18}$$

From the expression of $\mathfrak{L}^{-1}\{T(s)\}$ and of $g(s)$ we obtain, by a routine application of the procedures of the Laplace transformation:

$$G(t) = 1 - \sum_{p_n = p_n^+, p_n^-}\ \sum_{n=1}^{n=N} \frac{2\varphi}{1+\dfrac{\varphi_1\varphi_2}{p_n^2}}\left[\frac{\exp(s_n t)-1}{s_n} - \frac{\varphi_2}{\varphi_1+\varphi_2}\,\frac{\exp(s_n t) - \exp(-\varphi_1 t)}{p_n}\right],$$

where $p_n^+ = \varphi_1 + s_n^+$ and $p_n^- = \varphi_1 + s_n^-$ are the roots of the following equation in p_n:

$$p_n^2 - (\varphi_1 - \varphi_2 - \nu_n\xi)\, p_n - \varphi_1\varphi_2 = 0 . \tag{19}$$

From equation (19) we see that $p_n^+ > 0$, $p_n^- < 0$. Consequently, if we put

$$R_n = \sqrt{(\varphi_1 - \varphi_2 - \nu_n\xi)^2 + 4\varphi_1\varphi_2} ,$$

we easily obtain for $p_n = p_n^+, p_n^-$:

$$\frac{p_n^2 + \varphi_1\varphi_2}{|p_n|} = R_n .$$

Using this relation it may be seen that the term in $\exp(-\varphi_1 t)$ in $G(t)$ vanishes. Consequently, if we put

$$\epsilon_n^+ = \exp(s_n^+ t) , \qquad \epsilon_n^- = \exp(s_n^- t) ,$$

we can write:

$$G(t) = 1 + \sum_{n=1}^{n=N}\frac{2\varphi}{R_n}\left[\left(1+\frac{\varphi_1}{s_n^+}\right)(1-\epsilon_n^+) - \left(1+\frac{\varphi_1}{s_n^-}\right)(1-\epsilon_n^-) + \frac{\varphi_2}{\varphi_1+\varphi_2}(\epsilon_n^+ - \epsilon_n^-)\right].$$

The condition $G(0) = 1$ can be easily checked for this relation. The rela-

tion $G(\infty) = 0$ gives the following identity:

$$\sum_{n=1}^{n=N} \operatorname{cosec}^2 \frac{\theta_n}{2} = 2N^2. \tag{20}$$

If this is taken into account the expression of $G(t)$ can be further simplified:

$$G(t) = \frac{2\varphi\varphi_1}{\varphi_1 + \varphi_2} \sum_{n=1}^{n=N} \frac{\pi_n + \varphi_2 \sigma_n}{R_n}, \tag{21}$$

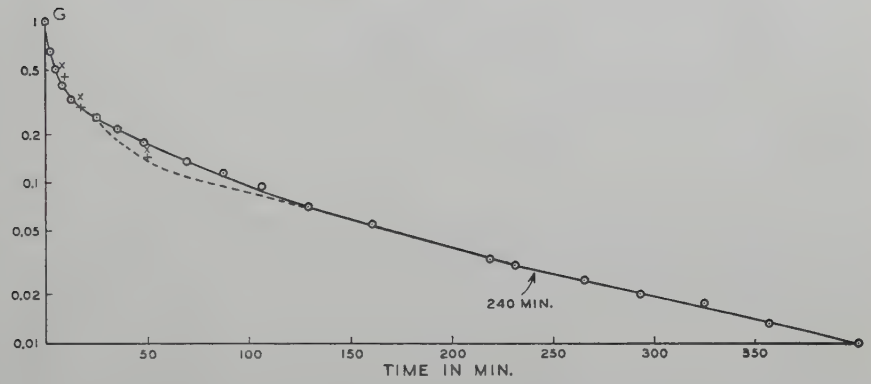

FIGURE 2. Analysis of diffusion of sodium out of a frog sartorius muscle (muscle N 46). Fraction of substance remaining in the tissue against duration of diffusion.

where

$$\pi_n = \frac{p_n^-}{s_n^-}\epsilon_n^- - \frac{p_n^+}{s_n^+}\epsilon_n^+, \qquad \sigma_n = \frac{\epsilon_n^-}{s_n^-} - \frac{\epsilon_n^+}{s_n^+}.$$

This is an expression for *the fraction of the initial amount of the substance which at time* t *exists in the tissue.*

3. *Determination of diffusion and permeability coefficients from experimental data.* Figure 2 represents the results of an analysis of a set of data kindly supplied by Drs. Harris and Burn. The circled points joined by a fully drawn curve and one surrounded by a square are experimental values on the diffusion of radioactive sodium from a frog sartorius muscle immersed in Ringer's solution. The thickness of the muscle was 800μ; its mass, 75 mg. The ordinate represents the fraction of radioactive sodium remaining in the muscle after immersion plotted on a logarithmic scale. The abscissa gives time in minutes. This set of data was analyzed with the objective of determining the diffusion coefficient D of sodium in the intercellular spaces and the permeability coefficients k_i and k_c of the cell membrane. It is apparent from the diagram that the function $G(t)$ tends

asymptotically to be exponential and this fact has been found very convenient for the calculation of the constants φ_1 and φ_2 which determine the permeability coefficients. Since the expression of $G(t)$ is of the type $\Sigma A_n \exp(s_n t)$ with A_n independent of t, the simple exponential shape of $G(t)$ for large t is due to the fact that from a certain t on the term containing the largest s_n is much larger than the sum of all the remaining terms. To analyze this fact it is necessary to examine the location of the quantities s_n or p_n which are the roots of equations (17) and (19) respectively. From equation (19) we see that p_n^+ and p_n^- have opposite signs; consequently, $p_n^+ > 0$, $p_n^- < 0$. Therefore $s_n^+ > -\varphi_1$, $s_n^- < -\varphi_1$. Since s_n^+ has the same sign as s_n^- we can see, using equation (17), that both s_n^+ and s_n^- are negative and can be represented diagrammatically as in Figure 3. Differentiating equation (19) with respect to n and taking into ac-

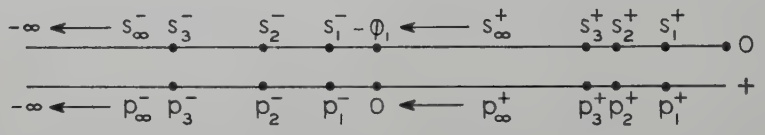

FIGURE 3

count the expression of ν_n we obtain:

$$\frac{R_n d|p_n|}{dn} = -\frac{2\pi}{N}\xi p_n \sin\theta_n .$$

Consequently $d|p_n^+|/dn < 0$ and $d|p_n^-|/dn > 0$. Therefore, as n increases p_n^+ approaches 0 and p_n^- moves away from zero, as indicated in Figure 3. Since the largest n is N and $\theta_N \to \pi$ as $N \to \infty$ it can be easily seen from equation (19) that $p_N^+ \to 0$, $p_N^- \to -\infty$.

For large t the terms in $G(t)$ containing $\exp(s_n^- t)$ are negligible with respect to those containing $\exp(s_n^+ t)$. A preliminary analysis of the data of Figure 2 showed that φ_2 was of the order of magnitude of 10^{-5} sec^{-1} whereas p_n^+ was between 10^{-6} and 10^{-8} sec^{-1}. Consequently, as a first approximation, p_n^+ could be neglected with respect to φ_2, and we put $\pi_n \approx 0$ in the expression of $G(t)$. In addition to this, the same analysis showed that $\varphi_1\varphi_2$ was of the order of magnitude of 10^{-9} sec^{-2} whereas $\nu_n\xi - (\varphi_1 - \varphi_2)$ was of the order of magnitude of 10^{-2} or 10^{-3} sec^{-1}. Consequently

$$R_n \approx |\nu_n\xi - (\varphi_1 - \varphi_2)| .$$

But, whereas $\nu_n\xi$ appeared to be of the order of magnitude of 10^{-2} or

10^{-3} sec^{-1}, $\varphi_1 - \varphi_2 \approx \varphi_1$ was of the order of 10^{-4}. Therefore:

$$R_n \approx \nu_n \xi = 4N^2 \varphi \sin^2 \frac{\theta_n}{2}.$$

Since p_n^+ was of the order of 10^{-6} to 10^{-8} sec^{-1} we put $s_n^+ \approx -\varphi_1$. In this way we obtained from equations (21) and (20), for large t,

$$G(t) \approx \frac{\exp(-\varphi_1 t)}{1 + \frac{\varphi_1}{\varphi_2}}. \tag{22}$$

This procedure has been found quite accurate in reference to the data of Figure 2. The curve between 130 and 240 min is practically a straight line which gives us, using formula (22),

$$\varphi_1 = 1.426 \times 10^{-4} \text{ sec}^{-1},$$

$$\varphi_2 = 3.724 \times 10^{-5} \text{ sec}^{-1}.$$

At $t = 240$ min the curve in Figure 2 seems to pass into another straight line which has been interpreted as a consequence of some change occurring in the muscle after the first four hours. This is analyzed in section 7. For the present the data of Figure 2 above $t = 240$ min may be disregarded.

The determination of φ is somewhat more laborious and requires an actual fitting of the curve $G(t)$ with its theoretical expression. A crude approximation to the value of φ can be obtained by comparing the theoretical expression of the slope of $G(t)$ at $t = 0$, which is

$$\frac{-2N\varphi\varphi_1}{\varphi_1 + \varphi_2}, \tag{23}$$

with the one obtainable from the experimental data. It will be shown shortly that 5 is a reasonable value to assign to N for the muscle in question. Estimating by numerical differentiation the slope of the graph at $t = 0$ in Figure 2 as -4.54×10^{-3} sec^{-1} and using the previously determined values of φ_1 and φ_2, formula (23) gives $\varphi \approx 5.7 \times 10^{-4}$ sec^{-1}. This has been found quite good but somewhat too small to fit the whole curve. A better value of φ has been obtained by using the exact expression of $G(t)$ and determining by trial and error that value of φ which gave the experimental value of G at $t = 25$ min. Two or three trials accompanied by an interpolation gave us:

$$\varphi = 7.4 \times 10^{-4} \text{ sec}^{-1}.$$

The plus signs in Figure 2 and the point surrounded by a square represent the values calculated with the above φ by means of the exact expression

(21) of $G(t)$. Beyond $t = 50$ min the theoretical curve smoothly approaches the experimental one.

Although for reasons of mathematical simplicity Figure 1 has been used to derive the equations of the theory, Figure 4 represents somewhat more closely the arrangement of the fibers in the muscle considered. Three fibers of circular cross-sections placed in two adjacent layers are represented in this figure. The centers of cross-sections of the fibers form an equilateral triangle. The minimum distance between two neighboring cells $2KW$ is somewhat smaller than the width $2W$ of a single intercellular diffusion

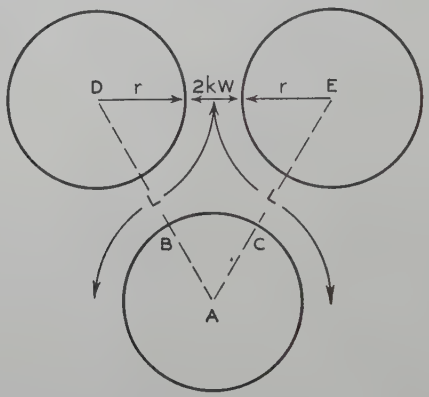

FIGURE 4. Schema of fiber arrangement in the muscle

flow, i.e., $K < 1$. The following relation holds between the radius r of the fiber, the thickness l of the tissue, and the number $2N$ of layers:

$$2N\,(r + KW)\,\sqrt{3} = l\,. \tag{24}$$

If we denote by η the fraction of the volume of the tissue occupied by the fibers, i.e., the ratio of the areas of three circular segments like ABC to the area of the triangle ADE, we obtain:

$$\pi r^2 = 2\,\sqrt{3}\,\eta\,(r + KW)^2\,.$$

Boyle et al. (1941) give for η a range from 0.87 to about 0.91. Mayeda (1890) gives for r an average value of 42μ. On the basis of these data the number $2N$ of layers in a tissue of $l = 800\mu$ can be calculated, since the previous two relations give:

$$(2N)^2 = \frac{2}{\pi\,\sqrt{3}}\,\frac{\eta l^2}{r^2}\,.$$

From this we find $2N = 10.77$ for $\eta = 0.87$. We take $2N = 10$ which cor-

responds to $r = 45\mu$ with the same value of η. This is still a reasonable magnitude for an average size of the fiber. The distribution of the radii of the sartorius muscle in frog shows, according to Mayeda (*loc. cit.*), a sharp peak around a mean of 42μ, but the full range goes from 10 to 76μ. The above value of r implies, using formula (24), $KW = 1\mu$. Accepting the values of φ_1 and φ_2 previously obtained and putting $L \approx \pi r, S \approx 2\pi r$, we obtain the coefficient of permeability:

$$k_c \approx \frac{\varphi_1 r}{2} = 3.2 \times 10^{-7} \text{ cm/sec},$$

TABLE I

Muscle Number	Weight Mg.	Thickness Cm.	φ_1 Sec.$^{-1}$	φ_2 Sec.$^{-1}$	k_c Cm. Sec.$^{-1}$	k_I Cm. Sec.$^{-1}$	k_c/k_I
N 46	75	0.08	1.43×10^{-4}	3.72×10^{-5}	3.21×10^{-7}	1.26×10^{-8}	25.5
N 67	75	0.08	3.30×10^{-4}	1.76×10^{-4}	7.43×10^{-7}	5.91×10^{-8}	12.6
N 69	90	0.1	5.32×10^{-4}	3.40×10^{-4}	12.0×10^{-7}	11.5×10^{-8}	10.5
N 70	90	0.1	7.11×10^{-4}	2.80×10^{-4}	16.0×10^{-7}	9.44×10^{-8}	17.0
N 71	85	0.08	3.87×10^{-4}	8.88×10^{-5}	8.72×10^{-7}	2.99×10^{-8}	29.2
N 72	85	0.08	4.11×10^{-4}	3.04×10^{-4}	9.25×10^{-7}	10.22×10^{-8}	9.06

All experiments done at 18° C.

which is about three times smaller than the values indicated by Harris and Burn, and

$$k_I \approx \varphi_2 r \frac{\eta^{-1} - 1}{2} \approx 1.26 \times 10^{-8} \text{ cm/sec}$$

which is about 10 times smaller than a value of Harris and Burn. In this way the ratio of sodium concentration outside of the cells, which is in equilibrium with the concentration inside, is ≈ 26 for the data used against a range of about 3.5 to 7 indicated by Harris and Burn. For the diffusion coefficient in the intercellular space we obtain:

$$D = 0.8\,(\pi N r)^2\,(\eta^{-1} - 1)\,\frac{r\varphi}{2W} \approx K \times 10^{-5} \text{ cm}^2/\text{sec}.$$

The coefficient 0.8 in this formula takes into account the solid substances existing in the intercellular spaces into which no sodium diffuses, K is the ratio of the minimum distance between two neighboring cells and the average width of the diffusion flow between these cells. Consequently $K < 1$ and D can be estimated as being about the order of magnitude of 10^{-6} cm^2/sec, which is in agreement with the calculations of Harris and Burn.

Table I gives results of calculations of the permeability coefficients

of a few other muscles, the data for which have also been kindly supplied by Drs. Harris and Burn. The muscle N 46 is the one whose analysis was just described in detail.

4. *Amount of substance inside and outside the cells.* If $Q_I(t)$ is the *amount of substance existing at* t *in all the intercellular spaces*, we have:

$$Q_I(t) = \sum_{n=1}^{n=N} Y_n(t).$$

The same amount expressed as a fraction of the amount initially existing in all the intercellular spaces is:

$$G_I(t) = \sum_{n=1}^{n=N} \frac{Y_n(t)}{Y}.$$

For the Laplace transform of the above we have:

$$g_I(s) = \mathfrak{L}\{G_I(t)\} = \frac{1-T}{s}.$$

Applying the same procedure as the one for $g(s)$ we obtain:

$$G_I(t) = 2\varphi \sum_{n=1}^{n=N} \frac{\pi_n}{R_n}. \tag{25}$$

Similarly for $Q_c(t)$, the *amount* of substance existing at t *in all the cells*, we find

$$Q_c(t) = \sum_{n=1}^{n=N} X_n(t).$$

If we express this as a fraction of the amount initially existing in all the cells we have:

$$G_c(t) = \varphi_1 \sum_{n=1}^{n=N} \frac{X_n(t)}{\varphi_2 Y}.$$

For the Laplace transform of this we have:

$$g_c(s) = \mathfrak{L}\{G_c(t)\} = \frac{1+\varphi_1 g_I(s)}{s+\varphi_1}.$$

From here we obtain:

$$G_c(t) = 2\varphi\varphi_1 \sum_{n=1}^{n=N} \frac{\sigma_n}{R_n}. \tag{26}$$

The following relation can be easily checked with the above equations:

$$G_I + \frac{\varphi_2}{\varphi_1} G_c = 1 + \left(\frac{\varphi_2}{\varphi_1}\right) G. \tag{27}$$

5. *Distribution of the substance among the various layers.* The amount of substance existing in single layers can be calculated without additional difficulty. From $y_m/(Y/N) = [1 - h_m(s)]/s$, using

$$h_m(s) = 2\,\frac{z^m + z^{2N+1-m}}{(1+z)(1+z^{2N})},$$

one can obtain the *amount of substance existing in the intercellular spaces of the* m*th layer*. Since

$$\mathfrak{L}^{-1}\{h_m(s)\} = 2N\varphi \sum_{n=1}^{n=N} |p_n| A_{m,n} \exp(s_n t)$$

where

$$A_{m,n} = \frac{\cos(m-1)\theta_n - \cos m\theta_n}{R_n},$$

we have

$$Y_m(t) = 2Y\varphi \sum_{n=1}^{n=N} A_{m,n}\pi_n. \tag{28}$$

The relation $Y_m(\infty) = 0$ gives us the identity:

$$\sum_{n=1}^{n=N} \frac{\sin(m-\frac{1}{2})\theta_n}{\sin\frac{\theta_n}{2}} = N.$$

The amount of substance *in the cells of the* m*th layer* can be calculated from the following equation which is easily obtainable from the relation between $C_n(t)$ and $I_n(t)$:

$$X_m(t)\exp(\varphi_1 t) = \frac{\varphi_2}{\varphi_1}\frac{Y}{N} + \varphi_2 \int_0^t Y_m(t)\exp(\varphi_1 t)\,dt.$$

We have:

$$X_m(t) = 2\varphi\varphi_2 Y \sum_{n=1}^{n=N} A_{m,n}\sigma_n. \tag{29}$$

Since for large t

$$\pi_n \approx 0, \qquad \sigma_n \approx \frac{\exp(-\varphi_1 t)}{\varphi_1},$$

as the diffusion goes on the amount of substance in the intercellular spaces becomes negligible with respect to that remaining inside the cells. Taking into account also that

$$R_n \approx 4N^2\varphi \sin^2\frac{\theta_n}{2},$$

it is easy to show that the amount of substance remaining in the cells tends to be approximately the same in each layer as t becomes large.

6. *Formulae at the beginning of the diffusion process.* For small t we ob-

tain by power series expansion of ϵ_n:

$$\pi_n = \frac{R_n}{N^2 \nu \varphi} - R_n t + \ldots\ldots,$$

$$\sigma_n = \frac{R_n}{N^2 \nu_n \varphi \varphi_1} - \frac{R_n}{2} t^2 + \ldots\ldots.$$

Therefore, if we take into account the trigonometric identities previously derived, we obtain for *the amount of substance existing in the whole tissue at the time* t, *divided by the initial amount:*

$$G(t) = 1 - \frac{2N\varphi\varphi_1}{\varphi_1 + \varphi_2} t + \ldots\ldots. \tag{30}$$

This amount is *distributed* as follows *among the intercellular spaces and the cells:*

$$G_I(t) = 1 - 2\varphi N t + \ldots\ldots, \tag{31}$$

$$G_c(t) = 1 - N\varphi\varphi_1 t^2. \tag{32}$$

Comparing these three relations with each other and keeping equation (27) in mind, it is seen that φ_1 and φ_2 appear in the first order term of $G(t)$ only because at $t = 0$ the distribution of the substance among the intercellular spaces and the cells depends on φ_1 and φ_2. At the very beginning the rate of diffusion out of the tissue does not depend on the permeability coefficients of the cells.

7. *Conditions at the boundary between the tissue and the solution.* An inspection of Figure 2 shows that the theoretical fit is not as good as one would desire for small t. It cannot be improved, however, since it would require a lowering of φ which would make the fit worse for larger t. The question arises whether the fit could be improved by taking into account the fact that the concentration of the substance at the boundary may not be so small as to be taken equal to zero, particularly at the beginning of the diffusion process. The objective of the present section is to analyze a possible effect of this situation.

Let us consider in the solution two layers adjacent to the tissue, and let us number them $n = 0$ and $n = 2N + 1$ (see Fig. 1). Let us assume that the concentration of the substance in the solution can be made equal to zero at points like P_0 within these layers not very distant from the boundary of the tissue. Consider a part of the layer $n = 0$ between two axes aa and bb and the diffusion flow between it and the adjacent intercellular space of the layer $n = 1$. We consider W_0 to be an average width of this diffusion flow (W_0 is somewhat larger than W, but smaller than V), D_0 to be the average diffusion coefficient of the substance considered in

the intercellular space and in the external solution. Then the number of molecules which diffuse per unit time within aabb from the intercellular space under consideration to the adjacent part of the solution (Fig. 1) is approximately:

$$\frac{D_0 W_0 I_1}{L_0} = \frac{H D W I_1}{L},$$

where L_0 is the average length of this diffusion flow, that is, about the length of an arc like P_1P_0. The coefficient H defined by the above equation is introduced for the sake of mathematical simplicity. Consequently, we find:

$$\frac{B d I_1}{dt} = \frac{2 D W}{L}(I_2 - I_1) - \frac{2 H D W}{L} I_1 - S(k_I I_1 - k_c C_1).$$

Since it is likely that D_0, W_0, L_0 are of the same order of magnitude as D, W, and L respectively, we will put $H = 1$ tentatively, which reduces the above differential equation to equation (2) taken for $n = 1$ with $I_0 = 0$. With this definition of I_0 and with the additional condition $I_{N+1} = I_N$ expressing the symmetry of the tissue with respect to a middle axis in the tissue, equation (2) can now be used for $n = 1, \ldots, N$, and equations (3) and (4) can be disregarded. Equation (2) is solved in a similar manner as before and $g(s)$ still has the same expression with the only difference that

$$T = \frac{z}{N} \frac{z^{2N} - 1}{(z-1)(z^{2N+1} + 1)}, \qquad \theta_n = \frac{(2n-1)\pi}{2N+1},$$

instead of equations (12) and (14). Equations (16) and (17) remain unchanged and the right-hand side of equation (18) has a factor $\gamma \cos^2(\theta_n/2)$ instead of 2φ, where $\gamma = 4\varphi N/(2N+1)$. The same replacement of 2φ has to be made in equations (21), (25), and (26). Equation (22) is unchanged, so that this theory yields the same permeability coefficients k_I and k_c as the previous one. The theoretical slope of the curve $G(t)$ at $t = 0$ is now only $\frac{1}{2}$ of the previous one. Equations (28) and (29) are now changed by the replacement of 2φ by γ and by a new meaning of $A_{m,n}$:

$$A_{m,n} = \frac{\sin\theta_n \sin(m\theta_n)}{R_n}.$$

The trigonometric identities which hold for the present angles θ_n are:

$$\sum_{n=1}^{n=N} \cot^2 \frac{\theta_n}{2} = (2N+1)N,$$

instead of equation (20), and

$$2\sum_{n=1}^{n=N} \sin(m\,\theta_n) \cot\frac{\theta_n}{2} = 2N + 1\,.$$

The coefficients of t and t^2 in equations (30) to (32) are now only one-half as large as before.

A fit of the experimental data of Figure 2 by means of the present theory has been done in the same way as before. The points marked by multiplication signs in Figure 2 are points of the theoretical curve which goes through the experimental point at $t = 25$ min, has the same values of φ_1 and φ_2 as before and $\varphi = 7.85 \times 10^{-4}$ sec^{-1}. It is seen that although the new fit is somewhat better than the previous one for larger t's, it is somewhat worse for smaller t's. The difference between the two values of φ is about 5%, which is completely irrelevant in view of the fact that φ is used only for the determination of the diffusion coefficient in the intercellular space, and this determination is affected by other uncertainties anyway. An interesting fact is that both theories lead to exactly the same permeability coefficients of the cell. Thus it does not seem possible to explain in this way the slight divergence of the theory from the experimental data of Figure 2, particularly at small t. However, the difference between the experimental data and the theory for small t may be partially accounted for by diffusion from the muscle in several ways not taken into account by the theory. We have considered the muscle to be an infinite sheet, whereas it actually has a small roughly triangular area instead. Thus diffusion actually occurs through all the edges of the muscle. Since approximately 10% of the total volume of the muscle lay within edges at a distance equal to twice the thickness of the muscle, the more rapid diffusion measured than given by theory is not surprising.

The dotted curve of Figure 2 represents a fit of the experimental data on muscle N 46 made by using the data above, $t = 240$ min, for the determination of φ_1 and φ_2 ($\varphi_1 = 1.12 \times 10^{-4}$ sec^{-1}, $\varphi_2 = 1.86 \times 10^{-5}$ sec^{-1}). The curve, calculated on the basis of the theory of this section, has been made to pass through the experimental point at $t = 25$ min which required $\varphi = 6.16 \times 10^{-4}$ sec^{-1}. It is seen that although the fit is very good above $t = 130$ min, the divergence is very strong for lower t's. As previously mentioned, this is the reason why the data above $t = 240$ min have been disregarded in final analysis.

This work was aided by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Boyle, P. J., E. J. Conway, F. Kane, and H. L. O'Reilly. 1941. "Volume of Interfibre Spaces in Frog Muscle and the Calculation of Concentrations in the Fibre Water." *Jour. Physiol.*, **99**, 401–14.

Churchill, R. V. 1944. *Modern Operational Mathematics in Engineering*. New York: McGraw-Hill.

Doetsch, G. 1937. *Theorie und Anwendung der Laplace-Transformation*. Berlin: J. Springer.

Harris, E. J. and G. P. Burn. 1949. "The transfer of sodium and potassium ions between muscle and the surrounding medium." *Trans. Faraday Soc.*, **45**, 508–28.

Mayeda, R. 1890. "Über die Kaliberverhältnisse der quergestreiften Muskelfasern." *Z. Biol.*, **27**, 119–52.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

SOME ELEMENTARY CONSIDERATIONS OF NEURAL MODELS

A. SHIMBEL
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

The outputs of nervous systems (as expressed in motor activity) are viewed as mathematical transformations on the inputs which enter via the sensory nerves. Simple nerve-ganglion models are exhibited which theoretically account for the arithmetic computations necessary to expedite such transformations.

Introduction. As a further development of the statistical approach to the theory of the central nervous system, it would be desirable to develop equations that would in some way delineate the complexity and versatility of the neural models implied by this orientation.

Simple nerve ganglion systems, for example, imply a variety of input-output problems. The output of such a system is in general a transformation of the input which is mediated by the intervening ganglion. It would be interesting to know what kinds of transformations are possible with ganglia whose only variables are size, distribution of synaptic delays, distribution of refractory periods, and distribution of periods of latent addition.

It may be that even such simple systems can exhibit various types of abstraction. The output (varying with time) of certain ganglia may, for example, give invariant responses to certain classes of time variable inputs.

Arithmetic Computation. If we assume that the function of the brain is to perform the proper mathematical operations (transformations) on the multi-modal inputs which reach it via the sensory nerves, and if we further assume that the arithmetic computations necessary to expedite such transformations are performed by simple neural arrays such as nerve-ganglion systems, it is necessary to indicate, at least theoretically, how such computation may be performed.

Addition. We must show, to begin with, that the nervous system is capable of responding to the sum of several stimuli of the same modality. A "possible" way in which the activity (average firing frequencies) of separate axone bundles can be summed is illustrated in Figure 1.

Each of the fibers of bundle number 1 enters the ganglion and synapses upon only one cell body. The same is true of bundle number 2. Each cell body of the ganglion gives off only one axone. These axones emerge from the ganglion to form the outgoing bundle (number 3).

Let f_i be the average firing frequency of a typical fiber in the ith bundle (average number of action currents per unit time *per axone*). Also let F_i denote the average firing frequency of the whole bundle.

By definition we may write $nf_i = F_i$, where n is the total number of fibers in the bundle. The model in Figure 1 is simply a composite of n units such as that illustrated in Figure 2.

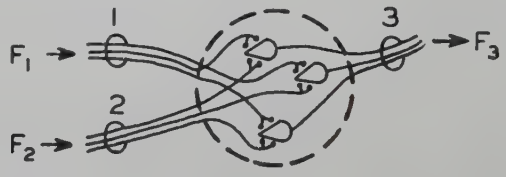

FIGURE 1

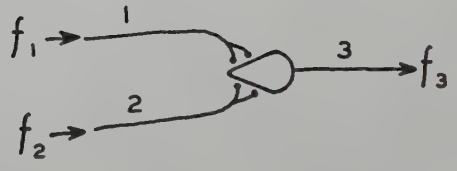

FIGURE 2

If we assume that either fiber 1 or fiber 2 is sufficient to fire fiber 3, then for low input frequencies $f_3 \approx f_1 + f_2$.

If we assume that f_1 and f_2 are Poisson-distributed then the probabilities that fiber number 1 or number 2 will fire within a δ of time (duration of refractoriness) are roughly δf_1 and δf_2 respectively for a small δ. The probability, therefore, that fiber number 3 will fire within the same period is given by

$$\delta f_3 = \delta f_1 + \delta f_2 - \delta^2 f_1 f_2 , \tag{1}$$

and therefore

$$f_3 = f_1 + f_2 - \delta f_1 f_2 . \tag{2}$$

But since $nf_i = F_i$ it follows from equation (2) that

$$F_3 = F_1 + F_2 - \frac{\delta}{n} F_1 F_2 . \tag{3}$$

Equation (3) tells us that F_3 is somewhat less than the sum of the inputs. The per-cent of error E is given by the expression

$$E = \frac{\frac{\delta}{n} F_1 F_2}{F_1 + F_2} \times 100 = \frac{\delta f_1 f_2}{f_1 + f_2} \times 100 . \tag{4}$$

Since the maximum error occurs when $f_1 = f_2 = f$, we find

$$(\max)\ E = \frac{\delta f^2}{2f} \times 100 = \frac{\delta}{2} f \times 10^2 . \tag{5}$$

Now if we let $\delta = 10^{-3}$ sec. and $f = 200$ impulses per second we get (max) $E = 10\%$ error. However, fibers rarely maintain frequencies of 200 impulses per sec. for very long. If we let f be somewhat more typical, say 25 impulses per second, then (max) $E = 1.25\%$ error.

For normal ranges of f_1 and f_2 we can therefore expect the mechanism in Figure 1 to record the sum of F_1 and F_2 to within less than 2% error.

Multiplication. The product of two input frequencies can be recorded by essentially the same mechanism as that illustrated in Figure 1. In this case, however, the thresholds of the ganglion cells are assumed to be higher so that both of the incoming fibers must fire (within the time of latent addition σ) in order that the outgoing axone should fire.

By similar considerations we find that roughly

$$F_3 = \frac{\sigma}{n} F_1 F_2 . \tag{6}$$

The output of such a system is proportional to the product of the inputs. Such a model, however, may not be too stable in view of the fact that the proportionality constant σ/n is so small. This implies that the output, although proportional to the product of the inputs, is relatively much smaller. The behavior of the system, particularly for large n, may become erratic due to random fluctuations in the ganglion.

A more stable model for recording the product of two frequencies would be possible if a ganglion could be constructed such that its output is proportional to the logarithm of its input. Actually all we would need for practical purposes is a ganglion whose output would be approximately proportional to the logarithm of its input for a substantial range of input frequencies.

Such a ganglion would make possible the construction of the model illustrated in Figure 3.

Subtraction. The mechanism illustrated in Figure 4 is designed to record the difference between two stimuli.

Bundle A sends axones to nucleus N_1 and also sends *inhibitory* branches to N_2. Bundles N_1 and N_2 give off axone bundles C and D respectively and also the combined branch bundle E. The activity in A tends to produce an activity in N_1 proportional to the activity in A. The inhibitory branches of B tend to inhibit the activity in N_1 proportional to the difference between the activities in A and B, so long as (a) is greater than (b). When (A) is smaller than (B) the activity in C is zero. Similarly the activity in D is proportional to the difference (B)—(A).

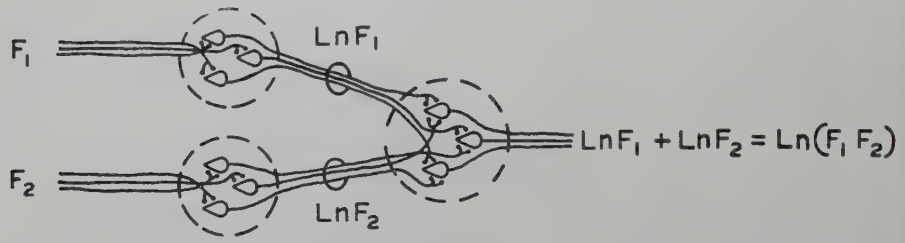

FIGURE 3

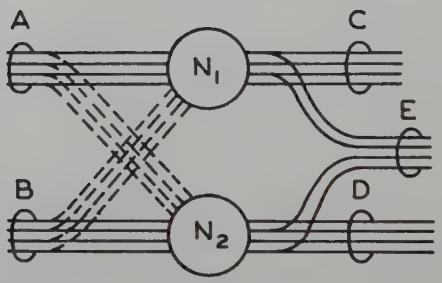

FIGURE 4

Finally, the activity in E is proportional to the absolute value of the difference between (A) and (B).

Ratio. The chapter on Discrimination of Relations in N. Rashevsky's book *Mathematical Biophysics* (1938) shows a simple mechanism which records the ratio of two stimuli. Many simple alternatives to this model are possible. Only a slight change of nomenclature and argument easily transforms Rashevsky's model to one consistent with the statistical approach.

Differentiation. Models designed to answer more complicated needs may require similar or identical information to be "fed" into two or more sub-systems at the same time. As has already been indicated in

Figure 4, such duplication of information can be achieved by the simple branching of the axones in the bundle under consideration. The arrival of information at its central destination may be delayed (if some mechanism requires such a delay) by having it pass first through one or more ganglia. At each such "station" the information will be delayed by about half of one millisecond.

Now with the above it is possible to construct a mechanism which will respond to the first derivative with respect to time of the activity in any input bundle. This is achieved by first duplicating the information (with a time lag as discussed above) and sending the two identical streams of information (but out of phase) into the mechanism illustrated in Figure 4. The resulting transformation would give the *difference* between the activity in the bundle at a time t and at a certain *constant* time earlier. But such a difference is roughly proportional to the first time derivative of the input as required.

Repetition of this device could be used for abstracting higher time derivatives of the same input.

Although models such as those discussed in the foregoing are interesting as such, it would be much more desirable to derive mathematical expressions, "existence theorems" as it were, which would tell us in a general way just how much is possible in the way of models starting with our initial assumptions.

This investigation is part of the work done under Contract No. AF 19(122)-161 between the U.S. Air Force Cambridge Research Laboratories and The University of Chicago.

LITERATURE

Shimbel, Alfonso. 1949. "Input-Output Problems in Simple Nerve-Ganglion Systems." *Bull. Math. Biophysics*, **11**, 165–71.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

INPUT OUTPUT CURVES OF AGGREGATES OF "SIMPLE COUNTER" NEURONS

ANATOL RAPOPORT
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

The refractory periods of an aggregate of simple "counter" neurons are assumed distributed according to some probability frequency. The output of the aggregate is computed for rectangular and triangular distributions. In particular, it is shown that the maximum output of an aggregate with any triangular distribution cannot exceed the maximum output of its average neuron by a factor greater than 2 ln 2. This puts an upper bound on the amount of departure from the behavior of the average neuron which an aggregate characterized by a certain type of distribution can show.

Next, the aggregate is supposed to be subjected to regularly spaced stimuli. Under these conditions, a single neuron will give a discontinuous output curve. If, however, the refractory periods are distributed according to some frequency, the output curve may be "smoothed out." A general condition on the distribution is derived which makes the output monotone increasing with the input. The condition is applied to some special cases.

If each neuron in an aggregate were characterized by the same parameters, and if the neurons did not interact with each other, then the output intensity of the aggregate expressed in firings per unit time per neuron would be identical with the output of each of its neurons. If, however, the parameters are distributed in accordance with certain probability frequencies, the output of the aggregate will be an average with respect to such distributions. For large aggregates, such an average can be expressed as an integral. For example, let the frequency distributions of the refractory period δ, the period of latent addition σ, and the threshold h (cf. Rapoport, 1950b) be given by $D(\delta)$, $S(\sigma)$, and $H(h)$ respectively. Then, if $y(x, \delta, \sigma, h)$ is the output function of a neuron with parameters δ, σ, and h, and subjected to the input x, and if these parameters are independently distributed, we shall have for the output intensity of the aggregate (firings per unit time per neuron):

$$Y(x) = \int_0^\infty \int_0^\infty \int_0^\infty y(x, \delta, \sigma, h)\, D(\delta)\, S(\sigma)\, H(h)\, d\delta d\sigma dh . \quad (1)$$

If the function y and the frequency distributions D, S, and H are known, Y can be computed. It is more likely, however, that experimental procedure may allow us to determine $Y(x)$. In that case, although the functions of the integrand are by no means uniquely determined, nevertheless some conclusions can be drawn concerning them on the basis of certain assumptions. If, for example, certain forms are assumed for three of the functions under the integrals, the fourth may be determined as a solution of an integral equation, provided certain conditions which guarantee the existence of the solution are satisfied.

Thus, formally we are dealing with two classes of problems, the direct problems, where the output function of the individual neuron and the parameter distributions are given, and the output of the aggregate is to be deduced; and the inverse problems, where the output of the aggregate plus some of the integrand functions are given, and the remaining integrand functions are to be deduced.

In the present paper we shall solve some very special cases of the direct problem, not so much for their own importance as in an attempt to gain insight into some more general aspects.

Let each neuron of an aggregate be a "simple counter," i.e., it responds to every stimulus it receives, provided the stimulus does not fall within its refractory period δ. Such neurons were discussed in earlier papers (e.g., Rapoport, 1950a). Then the individual output function is

$$y(x, \delta) = x(\delta x + 1)^{-1}, \tag{2}$$

and the aggregate output density will be given by

$$Y(x) = \int_0^\infty \frac{x D(\delta)}{\delta x + 1} \, d\delta. \tag{3}$$

For very large frequencies, we shall have

$$Y(\infty) = \int_0^\infty \frac{D(\delta)}{\delta} \, d\delta. \tag{4}$$

We shall suppose that $D(\delta)$ is continuous to the right throughout, i.e.,

$$\lim_{\delta \to \delta_0 + 0} D(\delta) = D(\delta_0). \tag{5}$$

Since $D(\delta)$ is a probability frequency distribution, we must have

$$\int_0^\infty D(\delta) \, d\delta = 1.$$

Hence the only questions concerning the convergence of the integral (4) arise in connection with the possible discontinuities of $D(\delta)/\delta$ at the

origin. Such questions are not of interest here. We may, however, for completeness of discussion, assign the value ∞ to $Y(\infty)$ if (4) diverges. Such may be the case, for example, if our aggregate contains neurons with arbitrarily small refractory periods (e.g., if $\lim_{\delta \to 0} D(0) \neq 0$).

Let us examine the behavior of $Y(x)$ for very large x for two very simple distributions of δ, namely, the rectangular and the triangular.

For a rectangular distribution, we have

$$\begin{aligned} D(\delta) &= 0, \text{ for } 0 \leqq \delta < \Delta - a ; \\ D(\delta) &= (2a)^{-1}, \text{ for } \Delta - a \leqq \delta < \Delta + a ; \\ D(\delta) &= 0, \text{ for } \delta \geqq \Delta + a . \end{aligned} \tag{6}$$

Here Δ is the mean value of δ.

In this case, equation (3) becomes

$$Y(x) = \frac{1}{2a} \int_{\Delta - a}^{\Delta + a} \frac{x \, d\delta}{\delta x + 1} = \frac{1}{2a} \ln \frac{(\Delta + a)x + 1}{(\Delta - a)x + 1}, \tag{7}$$

and

$$Y(\infty) = \frac{1}{2a} \ln \frac{\Delta + a}{\Delta - a}. \tag{8}$$

As a approaches zero, i.e., if the values of δ are strongly concentrated around their mean value, $Y(\infty)$ approaches $1/\delta$, as, of course, should be the case. As a approaches Δ, i.e., the rectangular distribution attains its maximum spread about the mean value, $Y(\infty)$ increases without bound, since the integral (4) must diverge in that case. Equation (8) may be conveniently represented by a series of the form

$$\begin{aligned} Y(\infty) &= \frac{1}{2a}\left[\frac{2a}{\Delta} + \frac{2}{3}\left(\frac{a}{\Delta}\right)^3 + \frac{2}{5}\left(\frac{a}{\Delta}\right)^5 + \ldots\right] \\ &= \frac{1}{\Delta}\left[1 + \frac{1}{3}\left(\frac{a}{\Delta}\right)^2 + \frac{1}{5}\left(\frac{a}{\Delta}\right)^4 \ldots\right], \end{aligned} \tag{9}$$

from which it appears that $Y(\infty)$ grows monotonically without bound at an increasing rate as the spread (a/Δ) of the distribution increases to unity.

Consider now the triangular distribution given by

$$\begin{aligned} D(\delta) &= 0, \text{ for } \delta < \Delta - a ; \\ D(\delta) &= m(\delta - \Delta + a), \text{ for } \Delta - a \leqq \delta < \Delta ; \\ D(\delta) &= -m(\delta - \Delta - a), \text{ for } \Delta \leqq \delta < \Delta + a ; \\ D(\delta) &= 0, \text{ for } \delta \geqq \Delta + a . \end{aligned} \tag{10}$$

The normalization of $D(\delta)$ requires that $m = 1/a^2$. Therefore

$$Y(x) = \frac{1}{a^2}\left[\int_{\Delta-a}^{\Delta} \frac{(\delta-\Delta+a)\,x}{\delta x+1}\,d\delta + \int_{\Delta}^{\Delta+a} \frac{-(\delta-\Delta-a)\,x}{\delta x+1}\,d\delta\right]$$
$$= \frac{1}{a^2}\left[a \ln \frac{\Delta x + ax + 1}{\Delta x - ax + 1} + \left(\Delta + \frac{1}{x}\right) \ln \frac{(\Delta x+1)^2 - a^2x^2}{(\Delta x+1)^2}\right]. \tag{11}$$

By the application of L'Hospital's rule, it can be readily seen that

$$\lim_{\substack{a\to 0 \\ x\to\infty}} Y(x, a) = \frac{1}{\Delta}. \tag{12}$$

For the maximum spread, on the other hand, that is, for $a = \Delta$, we have

$$Y(x) = \frac{1}{\Delta^2}\left[\Delta \ln(2\Delta x+1) + \left(\Delta + \frac{1}{x}\right) \ln \frac{2\Delta x+1}{(\Delta x+1)^2}\right], \tag{13}$$

so that

$$\lim_{x\to\infty} Y(x) = \frac{2\ln(2)}{\Delta} \sim \frac{1.4}{\Delta}. \tag{14}$$

Thus the asymptotic frequency of an aggregate of simple counter neurons, whose refractory period is spread in a symmetric triangular distribution about the average, is modified by a factor of about 1.4 relatively to the asymptotic frequency of its average neuron if the triangular distribution has the maximum spread. We shall see that for any smaller spread, the correction factor is smaller, as is intuitively evident. In fact, we shall derive a series, giving the correction factor in terms of a and Δ.

Note that setting $x = \infty$ in equation (11) gives us the asymptotic frequency of the triangularly distributed aggregate,

$$Y(\infty) = \frac{1}{a^2}\left[\int_{\Delta-a}^{\Delta} \frac{(\delta-\Delta+a)}{\delta}\,d\delta + \int_{\Delta}^{\Delta+a} -\frac{(\delta-\Delta-a)}{\delta}\,d\delta\right]$$
$$= \frac{1}{a^2}\left[(\Delta+a)\ln(\Delta+a) + (\Delta-a)\ln(\Delta-a) - 2\Delta\ln\Delta\right]. \tag{15}$$

This may also be writtten as

$$\frac{1}{a^2}\left\{\Delta \ln\left(1-\frac{a^2}{\Delta^2}\right) + a\left[\ln\left(1+\frac{a}{\Delta}\right) - \ln\left(1-\frac{a}{\Delta}\right)\right]\right\}. \tag{16}$$

Since $a < \Delta$, we may expand the logarithms obtaining

$$\frac{1}{a^2}\left[\Delta\left(-\frac{a^2}{\Delta^2} - \frac{a^4}{2\Delta^4} - \ldots\right) + 2a\left(\frac{a}{\Delta} + \frac{a^3}{3\Delta^3} + \ldots\right)\right]; \tag{17}$$

and, upon combining terms,

$$Y(\infty) = \frac{1}{\Delta}\left(1 + \frac{a^2}{6\Delta^2} + \dots\right), \tag{18}$$

which is a series of positive terms only and hence monotone increasing with a.

Consider now any unimodal distribution $D(\delta)$ symmetric about its mode (= mean). Let the value of δ be more concentrated around the mean Δ than those of a triangular distribution with the same range and mean. Then we shall say that the distribution $D(\delta)$ is "sharper" than the corresponding triangular distribution. More precisely, if, when we superimpose the two distributions one over the other, the ordinates of the triangular distribution are greater at the extremes and less near the mean than those of the distribution in question, then we shall say that the dis-

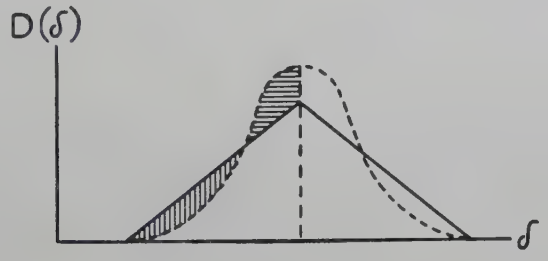

FIGURE 1

tribution in question is sharper than the triangular. Such a distribution is seen in Figure 1.

Since for symmetric distributions the area underneath the curve to the left of the mean is always equal to $\frac{1}{2}$, it follows that the two shaded areas in Figure 1 must be equal. But the contribution of the horizontally shaded area to the total output is greater than that of the vertically shaded area, because the neurons contributing to the former have smaller refractory periods. Hence on the left of the mean the total output of the triangular distribution is greater. The situation is reversed on the right side of the mean. However, the effect of the left side predominates, again because of the smaller refractory periods.

The asymptotic output of an aggregate with any symmetric distribution of refractory periods will always be greater than that of an aggregate all of whose neurons have a refractory period equal to the mean of the distribution. This again follows from the preponderance of the contribution to the aggregate output by the neurons whose refractory periods are smaller than the mean. Thus we have the result that in any symmetric

unimodal distribution sharper than the triangular one of the maximum possible spread ($a = \Delta$), the asymptotic frequency of the aggregate must lie between $1/\Delta$ and $2 \ln 2/\Delta \sim 1.4/\Delta$. For more accurate estimates, one may compare the distribution with a triangular one of the same range. The asymptotic output of the latter is given by the series (18), and the asymptotic output of an aggregate with a sharper distribution will not exceed it.

Thus the behavior of the input-output curves of aggregates with symmetric "bell-shaped" distributions (sharper than triangular ones) is not appreciably different from that of a single neuron with refractory period equal to the mean of the distribution. It can be readily verified that any such curve will have a derivative equal to unity at the origin and will approach its asymptotic value. Since the asymptotic values will not differ markedly from that of the single "average" neuron, we see that in general the input-output curves of most aggregates with quasi-normal distributions can be conveniently approximated by the input-output curve of the average representative neuron.

The Smoothing-out Effect of Distributions. If the stimuli received by the neurons in an aggregate are regularly spaced in time instead of Poisson-distributed, a somewhat different problem arises. It was shown earlier (Rapoport, 1950a) that the input-output curve of a single neuron subjected to regularly spaced stimuli is discontinuous, namely (cf. Rapoport, 1950a, Fig. 1),

$$\begin{aligned} y(x,\delta) &= x, \text{ for } 0 \leqq x < \frac{1}{\delta}; \\ y(x,\delta) &= \frac{x}{2}, \text{ for } \frac{1}{\delta} \leqq x < \frac{2}{\delta}; \\ &\vdots \\ y(x,\delta) &= \frac{x}{k}, \text{ for } \frac{k-1}{\delta} \leqq x < \frac{k}{\delta}. \end{aligned} \tag{19}$$

The output of an aggregate of such neurons, whose refractory periods have a distribution $D(\delta)$ will then be

$$Y(x) = \int_0^\infty y(x,\delta)\, D(\delta)\, d\delta. \tag{20}$$

We can now ask what will be the "smoothing-out" effect of the distribution function $D(\delta)$, in particular, what conditions must be satisfied by $D(\delta)$ in order that $Y(x)$ be monotone increasing with x?

As in our earlier paper (Rapoport, 1950a), we write $y(x, \delta)$ as a func-

tion of δ, holding x fixed, namely,

$$\begin{aligned} y(x,\delta) &= x, \text{ for } 0 \leqq \delta < \frac{1}{x}; \\ y(x,\delta) &= \frac{x}{2}, \text{ for } \frac{1}{x} \leqq \delta < \frac{2}{x}; \\ &\vdots \\ (x,\delta) &= \frac{x}{k}, \text{ for } \frac{k-1}{x} \leqq \delta < \frac{k}{x}. \end{aligned} \tag{21}$$

Then equation (20) becomes

$$Y(x) = x\left(\int_0^{1/x} D(\delta)\, d\delta + \frac{1}{2}\int_{1/x}^{2/x} D(\delta)\, d\delta + \ldots . \frac{1}{k}\int_{(k-1)/x}^{k/x} D(\delta)\, d\delta\right). \tag{22}$$

If $D(\delta)$ is continuous in δ, $Y(x)$ is a continuous function of x. We seek the condition on $D(\delta)$ such that $Y(x)$ is a *monotone increasing* function of x. We shall assume that $Y(x)$ possesses a derivative everywhere and ask under what conditions $Y'(x) \geq 0$ for all x.

Denoting $1/x$ by z, we re-write equation (22) as follows:

$$Z(z) = \frac{1}{z}\left(\int_0^{z} D(\delta)\, d\delta + \tfrac{1}{2}\int_z^{2z} D(\delta)\, d\delta + \ldots .\right). \tag{23}$$

The condition $Y'(x) \geq 0$ will now be replaced by the condition $Z'(z) \leqq 0$. Differentiating (23) with respect to z, we obtain

$$\begin{aligned} Z'(z) = -\frac{1}{z^2}\left(\int_0^{z} D(\delta)\, d\delta + \tfrac{1}{2}\int_z^{2z} D(\delta)\, d\delta + \ldots .\right) + \frac{1}{z}[D(z) \\ - D(0) + D(2z) - \tfrac{1}{2}D(z) + D(3z) - \tfrac{2}{3}D(2z) + \ldots .]. \end{aligned} \tag{24}$$

We shall assume that $Y(\infty) < \infty$ and hence that $D(0) = 0$. Then, collecting terms in (24), we obtain

$$\begin{aligned} Z'(z) = -\frac{1}{z^2}\left(\int_0^{z} D(\delta)\, d\delta + \tfrac{1}{2}\int_z^{2z} D(\delta)\, d\delta + \ldots .\right) \\ + \frac{1}{z}[\tfrac{1}{2}D(z) + \tfrac{1}{3}D(2z) + \tfrac{1}{4}D(3z) + \ldots .]. \end{aligned} \tag{25}$$

The expressions in both brackets are positive, so that to have $Z'(z) \leqq 0$

for all z, we must have

$$\left(\int_0^z D(\delta)\,d\delta + \tfrac{1}{2}\int_z^{2z} D(\delta)\,d\delta + \dots\right) \geqq z\left[\tfrac{1}{2}D(z) + \tfrac{1}{3}D(2z) + \dots\right], \tag{26}$$

for all z. Inequality (26) is the condition on $D(\delta)$ sought.

We shall investigate some special cases of $D(\delta)$ with regard to their "smoothing-out" effect on the aggregate output under the condition of regularly spaced stimulation. Consider, in particular, the rectangular distribution given by equations (6). Applying equation (22), we see that so long as x is sufficiently small (i.e., $1/x > \Delta + a$), $Y(x) = x$. Now let x increase and let us see what happens as $1/x$ just passes $(\Delta + a)$ from right to left, i.e., becomes less than that largest value of δ in the aggregate.

We note that $d(1/x) = -dx/x^2$. In particular, when $x = (\Delta + a)^{-1}$ $d(1/x) = -dx(\Delta + a)^2$. Since

$$\int_0^{\Delta+a} D(\delta)\,d\delta = 1,$$

the expression within the parentheses of the right side of (22) becomes

$$x\left(\int_0^{1/x} D(\delta)\,d\delta + \tfrac{1}{2}\int_{1/x}^{\Delta+a} D(\delta)\,d\delta\right). \tag{27}$$

(Note: We have broken off the series within the parentheses at the second term, because we are computing $Y(x)$ for those values of x for which $1/x \leq \Delta + a < 2/x$.) Let $x = (\Delta + a)^{-1} + dx$. The expression within the parentheses of (27) then becomes

$$1 - \frac{dx\,(\Delta + a)^2}{4a} \tag{28}$$

where the quantity subtracted from unity represents the loss of frequency due to the fact that the neurons for which the input frequency has just passed the critical frequency $(\Delta + a)^{-1}$ are responding to only every other stimulus. The total frequency per neuron for $x = (\Delta + a)^{-1} + dx$ is thus

$$[(\Delta + a)^{-1} + dx]\left(1 - \frac{dx\,(\Delta + a)^2}{4a}\right). \tag{29}$$

Neglecting higher order infinitesimals, we may write it

$$(\Delta + a)^{-1} + dx\left(1 - \frac{\Delta + a}{4a}\right). \tag{30}$$

Hence, at the critical input frequency $(\Delta + a)^{-1}$ the output curve will

remain increasing if, and only if,

$$\Delta + a < 4a; \qquad a > \frac{\Delta}{3}. \tag{31}$$

Inequality (31) gives the minimum spread of a rectangular distribution which will "smooth-out" (in the sense of introducing monotonicity) the output curve at the first discontinuity. The situation at the succeeding discontinuities can be similarly computed.

The derivative $Y'(x) = 1$ in the range $0 \leqq x < (\Delta + a)^{-1}$. The computation of $Y(x)$ at the points immediately to the right of $(\Delta + a)^{-1}$ gives

$$Y(x) = \frac{1}{4a} + x\,\frac{3a - \Delta}{4a}. \tag{32}$$

Therefore $Y'(x)$ in that region is given by

$$Y'(x) = \frac{3a - \Delta}{4a}. \tag{33}$$

We thus see that there will always be an abrupt change of slope in the output curve at the first critical point. The rectangular distribution cannot smooth out the output curve to make the derivative continuous. We shall see, however, that the triangular distribution can produce such a smoothing-out effect. For this case, if $\Delta + a > 1/x > \Delta$, equation (22) becomes

$$\begin{aligned} Y(x) &= x\left(\frac{1}{2} - \frac{1}{a^2}\int_{\Delta}^{1/x} (\delta - \Delta - a)\,d\delta - \frac{1}{2a^2}\int_{1/x}^{\Delta+a} (\delta - \Delta - a)\,d\delta\right) \\ &= x\left(\frac{3}{4} - \frac{1}{4a^2x^2} + \frac{\Delta + a}{2a^2x} - \frac{\Delta^2}{4a^2} - \frac{\Delta}{2a}\right). \end{aligned} \tag{34}$$

Then

$$Y'(x) = \frac{3}{4} + \frac{1}{4a^2x^2} - \frac{\Delta^2}{4a^2} - \frac{\Delta}{2a}. \tag{35}$$

Setting $x = (\Delta + a)^{-1}$, we obtain

$$Y'[(\Delta + a)^{-1}] = 1. \tag{36}$$

Thus not only is there no decrease in output initiated at the first critical frequency, but the output curve is continuous over it with a continuous derivative.

Nevertheless, $Y(x)$ may or may not be monotone if $D(\delta)$ is a triangular distribution. We shall make a sample investigation in the region where $(\Delta + a)^{-1} \leq x < \Delta^{-1}$ to determine under what conditions $Y'(x)$ can change sign in that interval, i.e., under what conditions $Y(x)$ will fail to be monotone. For this to happen, we evidently must have holding

simultaneously

$$\frac{3}{4}+\frac{1}{4a^2x^2}-\frac{\Delta^2}{4a^2}-\frac{\Delta}{2a}=0\,; \tag{37}$$

$$(\Delta+a)^{-1}\leqq x<\Delta^{-1}\,. \tag{38}$$

Moreover, since $D(0) = 0$, it follows that $\Delta > a$. Hence $\Delta + a < 2\Delta$. This condition insures that equation (27) describes $Y(x)$ in the entire range of x given by (38), since as $1/x$ approaches Δ from above, $2/x$ has not yet passed $\Delta + a$.

Solving (37) for x^2, we obtain

$$x^2=(\Delta^2+2a\Delta-3a^2)^{-1}\,. \tag{39}$$

Substituting this value into (38), we see that the following inequalities must be simultaneously satisfied:

$$\Delta^2+2a\Delta-3a^2\leqq(\Delta+a)^2\,; \tag{40}$$

$$\Delta^2+2a\Delta-3a^2>\Delta^2\,. \tag{41}$$

Inequality (40) is identically satisfied. Inequality (41) holds if, and only if,

$$a<2\,\frac{\Delta}{3}. \tag{42}$$

Thus inequality (42) is the condition that $Y(x)$ fails to be monotone in the interval $(\Delta + a)^{-1} \leq x < \Delta^{-1}$. It follows that if $Y(x)$ is to be monotone in that interval, we must have

$$a\geqq 2\,\frac{\Delta}{3}, \tag{43}$$

which gives the minimum spread of the triangular distribution insuring monotonicity in the region $(\Delta + a)^{-1} \leqq x < \Delta^{-1}$.

Proceeding in this way, we can examine the behavior of $Y(x)$ in succeeding intervals to see if more drastic conditions are required for its being monotone. One suspects that (31) is sufficient for the entire range in the case of the rectangular distribution, and, similarly, (43) is sufficient for the triangular distribution, since the satisfaction of these inequalities removes the *greatest* oscillation of the discontinuous function $y(x, \delta)$, namely, the first one (at $x = 1/\delta$). The proof of this conjecture is perhaps not difficult.

The general qualitative result one may deduce from the above discussion is that given a sufficient dispersal of δ (measured by a in the triangular distribution), the output curve of an aggregate of neurons subjected to regularly spaced stimuli can be smoothed out. Similar results are likely

to be obtained from more general distributions. In general the smoothing-out effect will depend on the dispersal of the distribution. However, the converse of this conclusion is also interesting. If the dispersal of a distribution is small, the aggregate may exhibit a smooth output curve in response to regularly spaced stimuli, and yet the curve may not be monotone (i.e., exhibit maxima). Here then is a theoretical possibility of having non-monotone output curves without postulating any inhibitory action whatsoever.

Now a regularly spaced sequence of stimuli can be considered as a distribution of intervals without dispersal. In general a time sequence of stimuli will not be absolutely regular, so that the distribution of the intervals between the stimuli will exhibit a finite dispersal. We have seen that the dispersal associated with a Poisson shower is sufficient to give a monotone output curve even of a single neuron (or an aggregate with zero dispersal of refractory periods). On the other hand, if the dispersal of the intervals between stimuli is zero, certain minimum dispersal of refractory periods is required for monotonicity. It follows that monotonicity can be obtained with some sort of combination of interval dispersal and refractory period dispersal. If this combined dispersal falls below a certain value, non-monotonicity appears. We have seen, however (Rapoport, 1950a), how a non-monotone output curve can be used to obtain sensitivity to particular ranges of frequency. It now appears that such an effect can be obtained by a "tightening up" of either the firing pattern or of the distribution of refractory periods or both.

This investigation is part of the work done under Contract No. AF 19(122)-161 between the U.S. Air Force Cambridge Research Laboratories and The University of Chicago.

LITERATURE

Rapoport, A. 1950a. "Contribution to the Probabilistic Theory of Neural Nets: I." *Bull. Math. Biophysics*, **12**, 109–21.

———. 1950b. "Contribution to the Probabilistic Theory of Neural Nets: II." *Ibid.*, **12**, 187–97.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

ON THE THEORY OF DIFFUSION OF ELECTROLYTES

I. OPATOWSKI
COMMITTEE ON MATHEMATICAL BIOLOGY
THE UNIVERSITY OF CHICAGO

In the theory of diffusion of electrolytes the following assumptions are frequently made: (i) the electrolytic solution is electrically neutral everywhere, (ii) the ionic concentrations and the electric potential all depend on a single Cartesian coordinate as the only space variable. Often the electric potential of the solution is determined on the basis of the Poisson equation alone, disregarding any other relation between this potential and the ionic concentrations. Since the Poisson equation only represents a condition which the potential fulfills, the use of this equation alone may lead to error unless the explicit relation for the potential involving a space integration of ionic concentrations is also taken into account. But if this relation is used the Poisson equation becomes redundant and, more important, assumptions (i) and (ii) appear unacceptable, the former because it leads to a zero electric potential everywhere, the latter because it is mathematically incorrect. The present paper is based on general equations of diffusion of ions, excluding the Poisson equation. These equations form a system of nonlinear integrodifferential equations whose number equals the number of ionic species present in the solution. It appears that when all ions are distributed symmetrically around a point all functions related to the above system of equations can be made dependent on a single space coordinate: the distance from the center of symmetry. Two methods of successive approximations are given for the solution of the equations in the case of spherical symmetry with limitation to the steady state. These methods are then applied to the study of the distribution of ionic concentrations and electrical potentials inside a cell of spherical shape in equilibrium with its surroundings. These methods are rapidly convergent; exact theoretical values of the electric potential are calculable on the boundary of the cell. It appears that the potential at the center of the cell is not more than $\sim$50% higher than at its boundary and that variation of concentration inside the cell is not very large. For instance, with 100 mV on the boundary the ionic concentration there is about four times higher than at the center. Calculations show that extremely small amounts of electricity are sufficient to account for the electric potentials currently observed. In a cell of 100 micra diameter an average concentration of only 10^{-14} mole/cm^3 of a monovalent ion would be sufficient to give 1 millivolt on the boundary. This concentration is directly proportional to the voltage and inversely proportional to the square of the cell diameter. Most of the numerical results given above are obtained by considering only those ions whose electrical charge is not compensated for by ions of an opposite sign. The total concentrations may be much higher than those quoted. The theory does not take into account possible effects of structural heterogeneities which may exist in the cell, particularly of various phase boundaries. An inciden-

tal result shows that the Boltzmann distribution function in the form employed in modern theory of electrolytes is fundamentally a consequence of the mathematical theory of diffusion alone. It is pointed out, however, that Boltzmann distribution is not always compatible with the definition of the electric potential.

I. *Some general equations of diffusion of ions.* The classical theory of diffusion of electrolytes is based on the differential equation

$$\frac{\partial c_i}{\partial t} = D_i \left[\text{divgrad } c_i + \frac{\epsilon_i}{kT} \text{div } (c_i \text{ grad } f) \right]. \tag{1}$$

This equation is an expression of the fact that ions move under the action of a concentration gradient grad c_i and of the electric potential f, which is the sum of an externally applied potential, if any, and of the potential

$$f(\mathrm{P}) = \frac{1}{K} \sum_i \epsilon_i \int_{V_\mathrm{Q}} \frac{c_i(\mathrm{Q})\, dV_\mathrm{Q}}{R(\mathrm{P}, \mathrm{Q})} \tag{2}$$

due to the ions themselves. In the above expressions $c_i(\mathrm{Q})$ is the concentration of the ions of kind i at Q; an infinitesimal volume element around this point being indicated by dV_Q; $f(\mathrm{P})$ is the potential at P, $R(\mathrm{P}, \mathrm{Q})$ is the distance between P and Q, the charge of the ion of kind i is indicated by ϵ_i, and the dielectric constant of the solution by K, the diffusion coefficient of the ith ion in the solvent is indicated by D_i, the interaction between the ions of various kinds is assumed to be of a purely electrical nature, that is, mutual diffusion effects are neglected. Electrical forces between the ions of the solute and the polarized molecules of the solvent are also neglected. The meaning of the other symbols is: t = time, k = Boltzmann constant, and T = absolute temperature. The mobility of the ions is taken to be $u_i = D_i/kT$, which implies an application of the law of perfect gases to the ions in solution.

We will limit ourselves in this note to the case in which no external electric field is acting on the solution. Under this condition equation (1) was used by M. Planck (1850) in his theory of electrolytic solutions. Instead of equation (2) he used the Poisson equation

$$\text{divgrad } f = -4\pi \sum_i \frac{\epsilon_i c_i}{K}, \tag{3}$$

but neglected its right-hand side. K. Sitte (1934, 1935) pointed out that this procedure is not justified because, using equation (2), it would be equivalent to putting electric potential equal to zero everywhere, which would reduce equation (1) to a diffusion equation of neutral molecules and not of ions. J. J. Hermans (1936) tried unsuccessfully to justify this approximation by a mathematical argument. Of course, a method could

be devised for the solution of the above equations by expanding the functions in MacLaurin series of $1/K$. Such a method has been used in extending the Debye-Hückel theory of electrolytic solutions (Gronwall, La Mer and Sandved, 1928; La Mer, Gronwall and Greiff, 1931). As in Planck's theory the zero order terms in such a method would imply a zero right-hand side in equation (3).

If the concentrations $c_i (i = 1, 2, \ldots, n)$ are considered as the unknowns, the functions c_i are solutions of n nonlinear integrodifferential equations obtainable by substituting f from equation (2) into (1). Instead of these equations Planck and Sitte worked with equations (1) and (3) and made the additional assumption that all the functions c_i and f depend on one Cartesian coordinate as the only space variable. However an inspection of equation (2) shows that the electric potential f and the density of the electric charges $\epsilon_i c_i$ to which f is due cannot all be functions of only a single Cartesian coordinate.

Since equations (1) and (2) are very difficult it is worthwhile to look for cases in which concentrations and electric potential can be made dependent on a single space coordinate. We will show that this is the case if the system has spherical symmetry. If the ionic concentrations c_i depend only on the distance r from a fixed point we have

$$f = \frac{4\pi}{K} \sum_i \epsilon_i \left(\frac{1}{r} \int_{r_1}^{r} r^2 c_i dr + \int_{r}^{r_2} r c_i dr \right), \tag{4}$$

where $r = r_1$ and $r = r_2$ are two bounding surfaces of the solution with $r_1 < r_2$. If $r_1 = 0$ the solution fills a sphere of radius r_2. Writing equation (1) in spherical coordinates we obtain:

$$\frac{\partial c_i}{\partial t} = \frac{D_i}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial c_i}{\partial r} \right) + \frac{\epsilon_i}{kT} \frac{\partial}{\partial r} \left(r^2 c_i \frac{\partial f}{\partial r} \right) \right], \tag{5}$$

and eliminating f we have

$$\left.\begin{aligned} \frac{\partial c_i}{\partial t} &= \frac{D_i}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c_i}{\partial r} \right) - D_i k_i \frac{\phi_i}{r^2} \\ \phi_i &= \frac{\partial}{\partial r} \left(c_i \sum_j \epsilon_j \int_{r_1}^{r} r^2 c_j dr \right) = \frac{1}{4\pi} \frac{\partial}{\partial r} (c_i Q) , \end{aligned}\right\} \tag{6}$$

where $k_i = 4\pi\epsilon_i / KkT$ and

$$Q = 4\pi \sum_i \epsilon_i \int_{r_1}^{r} r^2 c_i dr \tag{7}$$

is the total electric charge existing within a sphere of radius r and center

at the origin. Equation (6) can also be written as:

$$\frac{r^2}{D_i}\frac{\partial c_i}{\partial t}=\frac{\partial}{\partial r}\left(r^2\frac{\partial c_i}{\partial r}\right)-h_i\frac{\partial (c_iQ)}{\partial r}, \tag{8}$$

where $h_i = \epsilon_i/KkT$. It is seen from equation (7) that if the electric charge $Q(r)$ in a sphere $r = \text{const.}$ is a maximum or a minimum the surface $r = \text{const.}$ is electrically neutral. From equations (4) and (7) we have for the electrical potential:

$$f=\frac{Q}{Kr}+\frac{4\pi}{K}\sum_i \epsilon_i\int_r^{r_2} r\,c_i\,dr\,. \tag{9}$$

The above equations describe a very general type of diffusion phenomena in which c_i and f both are expressible in terms of a single space coordinate r.

II. *The steady state.* If sources and sinks of ions have appropriate asymptotic characteristics as $t \to \infty$ or if they cease to act as $t \to \infty$, a steady state is reached: $\lim \partial c_i/\partial t \to 0$ when $t \to \infty$, i.e., using equation (5) we have:

$$r^2\frac{dc_i}{dr}+\frac{\epsilon_i}{kT}r^2c_i\frac{df}{dr}\to A_i\,, \tag{10}$$

where the A_i are constants. The number of ith ions flowing per unit time at $t = \infty$ through a spherical surface $r = \text{const.}$ is $4\pi D_iA_i$. Equation (10) can be rewritten, using equation (8), in the form:

$$r^2\frac{dc_i}{dr}-h_ic_iQ\to A_i\,. \tag{11}$$

The steady state is ruled by n of these relations, if n is the number of ionic species. Eliminating Q between any two of these relations we see that any c_i can be expressed in terms of any other one:

$$[E_i(r^*,r)]^{\epsilon_j/\epsilon_i}c_j(r) = c_j(r^*) + A_j\int_{r^*}^{r}[E_i(r^*,r)]^{\epsilon_j/\epsilon_i}\frac{dr}{r^2}, \tag{12}$$

where r^* is a value of r at which all the concentrations are known a priori, and

$$E_i(r^*,r) = \frac{c_i(r^*)}{c_i(r)}\exp\left(A_i\int_{r^*}^{r}\frac{dr}{r^2c_i(r)}\right).$$

Thus the determination of the relation between the various ionic concentrations in steady state requires, in general, not only the knowledge of the amounts of ions flowing in steady state through a boundary (constants A_i), but also the knowledge of all the steady state concentrations $[c_i(r^*)]$ at least at one point ($r = r^*$) of the solution. In general these

data can be obtained through an analysis of the non-steady state, which we plan to discuss in a successive note. There are, however, important cases in which these constants can be secured without solving the non-steady state equations. We will consider one of these cases in the following section. In any event the system of n equations (11) can be changed by means of equations (7) and (12) into n mutually independent equations in the c_i's. These integrodifferential equations are very complicated however.

It can be seen from equation (11) that a *uniform distribution of ions in steady state is impossible unless the solution is everywhere neutral, that is, the electric potential is zero everywhere.* In fact the condition that c_i be independent of r implies $Q =$ constant, which is possible only if Q is identically zero.

From now on we will limit ourselves to the case in which there are *no sources and sinks of ions acting at* $t = \infty$. Then $A_i = 0$ and equation (10) gives

$$c_i \to c_{i(0)}\, e_i(f) \qquad \text{where} \qquad e_i(f) = \exp \frac{-\epsilon_i f}{kT} \tag{13}$$

and $c_{i(0)}$ are constants depending only on i. The above is the well known *Boltzmann equation* in the form currently used in the Debye-Hückel theory of electrolytes (see e.g. Harned and Owen, 1950, p. 25). It is interesting that this equation *can be derived from the mathematical theory of diffusion alone* which is based on no assumptions beyond the principle of mass conservation, Fick's law of diffusion, the formula $u_i = D_i/kT$ implying the law of perfect gases, and the assumption of direct proportionality of ionic velocities to electric forces. However the *Boltzmann equation is not always compatible with the definition of the electric potential.* In fact if the concentrations c_i are all functions of one Cartesian coordinate only, equation (13) requires that the potential f also be such a function. But this contradicts equation (2). Thus, to a thermodynamical inconsistency of the Boltzmann equation noted by L. Onsager (1933), a new one is herewith added.

If there is a point in the solution at which the potential is zero, the concentrations c_i at that point equal $c_{i(0)}$. If N_i is the total number of ions of kind i existing at $t = \infty$, then

$$4\pi\, c_{i(0)} \int_{r_1}^{r_2} r^2 e_i(f)\, dr = N_i . \tag{14}$$

From equation (13) we have the following relation between concentrations and electric potentials at any two points $r = a$ and $r = b$:

$$f(a) - f(b) = \frac{kT}{\epsilon_i} \ln \frac{c_i(b)}{c_i(a)} . \tag{15}$$

The determination of f in terms of r requires the solution of the integral equation in f:

$$\frac{K}{4\pi} f(r) = \sum_i \epsilon_i c_{i(0)} \left(\frac{1}{r} \int_{r_1}^{r} r^2 e_i(f)\, dr + \int_{r}^{r_2} r\, e_i(f)\, dr \right) \qquad (16)$$

which is obtainable from equations (4) and (13).

Equation (16) determines f as a function of r and of the $c_{i(0)}$'s. If N_i instead of $c_{i(0)}$ are known a priori, we can use the following equation instead of (16):

$$f(r) = \frac{1}{K} \sum_i \frac{\epsilon_i N_i}{\int_{r_1}^{r} r^2 e_i(f)\, dr} \left(\frac{1}{r} \int_{r_1}^{r} r^2 e_i(f)\, dr + \int_{r}^{r_2} r\, e_i(f)\, dr \right), \qquad (17)$$

which is easily obtainable from equations (14) and (16). Both equations (16) and (17) are convenient for a determination of $f(r)$ by successive approximations. The potential on the external boundary of the solution is easily calculable as follows:

$$f(r_2) = \frac{\sum_i e_i N_i}{K r_2}. \qquad (18)$$

This formula, which is also a simple direct consequence of the definition of the electrostatic potential, shows that the potential on the external boundary ($r = r_2$) is zero if the solution is totally neutral

$$\left(\sum_i \epsilon_i N_i = 0 \right).$$

For the potential of the internal boundary we have

$$f(r_1) = \frac{\sum_i \epsilon_i N_i \int_{r_1}^{r_2} r\, e_i(f)\, dr}{K \int_{r_1}^{r_2} r^2 e_i(f)\, dr}. \qquad (19)$$

III. *Solution of the equations of the steady state. Electric potential and ionic concentration in a spherical cell.*

A. *Method* 1. Consider as before a solution bounded by two spherical surfaces of radii $r = r_1$ and $r = r_2$, or a solution filling out a whole sphere of radius r_2. If f does not vary a great deal in the solution we can put $f \approx$ constant as a starting approximation in an iteration process based on equation (17). We then obtain an approximate expression $f_1(r)$ of the

potential $f(r)$ as follows:

$$f_1(r) = \frac{1}{K}\left(\frac{3r_2^2 - r^2}{2} - \frac{r_1^3}{r}\right)\frac{\sum_i \epsilon_i N_i}{r_2^3 - r_1^3}. \tag{20}$$

Since $|f_1(r)|$ is a decreasing function of r, thus—in a first approximation at least—the *electric potential has its maximum magnitude at the center and minimum on the boundary*. At $r = r_1$ we have

$$f_1(r_1) = \frac{3}{2K}\frac{(r_1 + r_2)\sum_i \epsilon_i N_i}{(r_1 + r_2)^2 - r_1 r_2}. \tag{21}$$

Therefore, the *steady state potential at the center of the sphere* ($r_1 = 0$) *equals, in first approximation, three halves of its value at the boundary*. By putting $f = f_1$ in the right-hand side of equation (17) a second approximation, f_2 of f, is obtained. It can be expressed in terms of tabulated functions if the solution fills a whole sphere, that is, if $r_1 = 0$,

$$f_2(r) = \frac{1}{K r_2}\sum_i \frac{\exp(H\epsilon_i r_2^2) - \frac{1}{r}\int_0^r \exp(H\epsilon_i r^2)\,dr}{\exp(H\epsilon_i r_2^2) - \frac{1}{r_2}\int_0^{r_2} \exp(H\epsilon_i r^2)\,dr}\epsilon_i N_i, \tag{22}$$

where

$$H = \frac{\sum_i \epsilon_i N_i}{2KkTr_2^3} = \frac{f(r_2)}{2kTr_2^2}.$$

[For a bibliography of the tables of the integral appearing in formula (22) see Fletcher, Miller and Rosenhead, 1946, p. 219.] Both approximations f_1 and f_2 give the *exact expression of the potential* f *at* $r = r_2$. At $r = 0$ the approximation f_2 gives

$$f_2(0) = \frac{1}{K r_2}\sum_i \frac{[\exp(H\epsilon_i r_2^2) - 1]\,\epsilon_i N_i}{\exp(H\epsilon_i r_2^2) - (r_2\sqrt{H\epsilon_i})^{-1}\int_0^{r_2\sqrt{H\epsilon_i}} \exp(x^2)\,dx}. \tag{23}$$

As an application let us consider a spherical cell of 100μ diameter in steady state, with no ionic sources inside, and in equilibrium with its surroundings. A part of the positive and negative ions existing in the cell neutralize themselves. Therefore they can be disregarded in studying the electrical properties of the cell; these are due only to the ions giving a net contribution to the density of the electric charge in the cell. For simplicity we assume that these latter ions are all of the same valence and call ϵ the

charge of each of them and N their total number in the cell. The net electric charge of the cell is then $N\epsilon$. Since the sign of the charge does not affect the formulae except for their sign, let us assume $N\epsilon > 0$ and the corresponding ions as monovalent. If we put $K = 80$ then formula (18) gives us the number of ions in question per each mV of electric potential on the boundary of the cell, as follows:

$$\frac{N}{f(r_2)} = \frac{K r_2}{300{,}000\epsilon} \approx 2{,}780 \text{ ions/mV}.$$

Thus *less than* 3,000 *ions are sufficient to produce* 1 *mV voltage on the boundary of a cell of* 100μ *diameter. This number of ions is directly proportional to the voltage and to the cell diameter.* Since the dielectric constant of a cell interior is likely to be smaller than 80, the number of ions required is also likely to be smaller than the above figure. In any event this is a *very small number of ions.* This fact appears particularly clear if we consider the *average ionic concentration in the cell* c_{av} *which gives* 1 *mV on its boundary* (monovalent ions):

$$c_{av} = 2.75 \times 10^{-21} \frac{K}{r_2^2} \text{ mole/cm}^3 \text{ mV}.$$

With $K = 80$ and $2r_2 = 100\mu$ this gives $c_{av} = 8.8 \times 10^{-15}$ mole/cm^3 mV, a quantity *inversely proportional to the square of the cell diameter.*

Both $Hr_2^2 = f(r_2)/2kT$ and $f_2(0)/f_1(0)$ depend on the potential $f(r_2)$ at the cell surface but not on its radius r_2. If $f_2(0) \approx f_1(0)$, $f_1(r)$ is a good approximation to $f(r)$. This fact is seen to depend on $f(r_2)$ but not on r_2.

The table below gives *the first and the second approximations to the voltage at the center of the cell* $f_1(0)$ *and* $f_2(0)$ *for various voltages* $f(r_2)$ *on the boundary of the cell at* T = 300° K. *The results do not depend on the cell diameter.*

$f(r_2)$ in	mV	1	50	100
$f_1(0)$ in	mV	1.5	75	150
$f_2(0)$ in	mV	1.5	69	128

It is seen that the approximation method for the calculation of f is quite good, particularly for lower voltages. For higher voltages the result can be improved by calculating a third approximation. This is obtained by putting $f = f_2$ in the right-hand side of equation (17) or (19) and carrying out numerically the necessary integrations. As the voltage $f(r_2)$ on the external surface of the cell tends to zero, $H\epsilon r_2^2 \to 0$ and equation (23) gives $f_2(0) \to 3f(r_2)/2$, in agreement with the first column of the above table.

B. *Method* 2. In the previous successive approximation method the starting approximation was a constant potential throughout the cell. Instead of the latter we can take a suitable step-function as the starting approximation. We illustrate this procedure by calculating the potential of the cell using the data of the previous example. As a starting approximation we put $f(r) = f(0)$ for $0 \leq r \leq r_2/2$ and $f(r) = f(r_2)$ for $r_2/2 < r \leq r_2$. Since $f(r_2)$ is directly calculable from equation (18), relation (19) gives a transcendental equation for $f(r_1)$:

$$\exp(x_1 - x_2) = \frac{3x_2 - x_1}{7x_1 - 9x_2},$$

where $x_i = \epsilon f(r_i)/kT$, with $r_1 = 0$. The solution of this equation for $r_2 = 100\mu$, $r_1 = 0$, $T = 300°$ K, $f(r_2) = 100$ mV gives a first approximation to the potential at the center $f_1(0) \approx 135$ mV, which is in quite good agreement with the results of the previous method. Equation (17) can now be used to determine $f_1(r)$, an approximation to the potential $f(r)$ in the whole range $0 \leq r \leq r_2$:

$$f_1(r) = \frac{\epsilon N}{K r_2^3} \frac{a_1 + a_2 \exp(x_1 - x_2)}{1 + 7 \exp(x_1 - x_2)}, \qquad 2\,(4)$$

where

$$a_1 = \begin{cases} 3r_2^2 - 4r^2 & \text{for} \quad r \leq \frac{r_2}{2} \\ \frac{r_2^3}{r} & \text{for} \quad r \geq \frac{r_2}{2} \end{cases}$$

$$a_2 = \begin{cases} 9r_2^2 & \text{for} \quad r \leq \frac{r_2}{2} \\ 12r_2^2 - \frac{r_2^3}{r} - 4r^2 & \text{for} \quad r \geq \frac{r_2}{2}. \end{cases}$$

If, following method 1, we take $x_1 = x_2$ formula (24) reduces to (20). A second approximation to $f(r)$ can be obtained by putting $f(r) = f_1(r)$ in the right-hand side of equation (17) and calculating the corresponding integrals. The ratio of local concentrations on the boundary and at the center is $c(r_2)/c(0) \approx 3.9$, according to formula (13), in the case of 100 mV on boundary, if we assume 135 mV as the correct figure for the potential at the center.

The following *modification* of method 2 has been suggested to the writer by Professor H. D. Landahl. Since, according to equation (17)

$$\left(\frac{\partial f}{\partial r}\right)_{r=r_1} = 0, \left(\frac{\partial f}{\partial r}\right)_{r=r_2} = -\frac{\sum_i \epsilon_i N_i}{K r_2^2},$$

we can use as the starting approximation to $f(r)$ an appropriate function having the calculated values $f(r_1)$ and $f(r_2)$ as well as the above slopes $\partial f/\partial r$ at $r = r_1$ and $r = r_2$. Such a function can be easily represented by a curve since $f(r)$ has no maxima or minima in $r_1 < r < r_2$ [cf. equation (17)]. This modification is likely to yield a more rapidly convergent result than method 2. It requires numerical integrations throughout, however, unless the curve can be fitted by a suitable analytical expression.

The distribution of potential in a spherical cell has been treated recently by F. Booth (1951) using the method mentioned above of Gronwall, La Mer and Sandved. Numerical computations of the potential using this method are laborious and the power series involved presents several disadvantages which have been discussed by Booth (*loc. cit.*).

This work was aided by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of The University of Chicago.

LITERATURE

Booth, F. 1951. "The solution of some potential problems in the theory of electrolytes." *Jour. Chem. Phys.*, **19**, 821–26.

Fletcher, A., J. C. P. Miller and L. Rosenhead. 1946. *An Index of Mathematical Tables*. New York: McGraw Hill.

Gronwall, T. H., V. K. La Mer and K. Sandved. 1928. "Über den Einfluss der sogenannten höheren Glieder in der Debye-Hückelschen Theorie der Lösungen starker Elektrolyte." *Physik. Zeitschr.*, **29**, 358–93.

Harned, H. S. and B. Owen. 1950. *Electrolytic Solutions*. 2nd Ed. New York: Reinhold Publishing Co.

Hermans, J. J. 1936. "Diffusionspotentiale und Ionenaktivitäten." *Zeitschr. physik. Chem.*, **176A**, 55–64.

La Mer, V. K., T. H. Gronwall and L. J. Greiff. 1931. "The influence of higher terms of the Debye-Hückel theory in the case of unsymmetric valence type electrolytes." *Jour. Phys. Chem.*, **35**, 2245–88.

Onsager, L. 1933. "Theories of Concentrated Electrolytes." *Chem. Rev.*, **13**, 73–89.

Planck, M. 1890. "Über die Erregung von Electricität und Wärme in Electrolyten." *Ann. Physik*, **39**, 161–86.

———. 1890. "Über die Potentialdifferenz zwischen zwei verdünnten Lösungen binärer Electrolyte." *Ibid.*, **40**, 561–76.

Sitte, K. 1934. "Untersuchungen über Diffusion in Flüssigkeiten." *Zeitschr. f. Physik*, **91**, 622–50.

———. 1935. "Zur Theorie der Diffusion von Elektrolyten." *Ibid.*, **93**, 698–701.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 14, 1952

PROLEGOMENA TO A DYNAMICS OF IDEOLOGIES

N. Rashevsky
Committee on Mathematical Biology
The University of Chicago

The paper outlines a possible further development of the suggestion made by the author in his recent book. Ideology is defined here as a verbalized, or at least verbalizable, behavior pattern which may be adopted by society. After a brief discussion of possible classifications of ideologies, a study is made of those ideologies which refer to the question of social and ethical interrelations. Different kinds of ideologies may be represented by *behavior matrices*, introduced in the author's book. A uniparametric representation of all such matrices is suggested and discussed. Next the previous results on social imitation are extended to the case of n different behaviors, each of which is determined by a particular value of a continuously varying parameter. It is shown that, depending on some other social parameters, changes from one ideology to another may proceed either quasi-continuously or definitely discontinuously. The paper concludes with some general speculations on the possibility of applying the above results to a mathematical interpretation of history.

In our recent book (Rashevsky, 1951a) we suggested the possibility of developing a quantitative dynamics of ideas, considering ideas as complex verbal behavior patterns. From a neurobiological point of view the overt behavior and the covert idea of the behavior merely form two parts of the same neurobiophysical process. The word "idea" was understood as meaning merely a verbalized, or at least verbalizable, statement of the process of behavior. It has been shown by E. Jacobson (1931) that when an individual thinks of reciting a given poem action currents are produced in the muscles of his tongue which show a time pattern identical with that of the action current observed when the poem is actually recited. However, the intensity of the action current in the latter case is about 10^3 times greater. This and other experiments by Jacobson (1930) make it plausible to assume that the thought or idea of a behavior is the result of a neurobiological process which forms an essential part of the neurobiological processes that result in the behavior itself. Thus no metaphysical meaning whatsoever is attached here to the word "idea."

However, since this word has to many persons a definite connotation, different from the one used here (e.g., Platonic ideas), we shall use instead

the word "ideology," which, though perhaps not as appropriate, is free from the connotation mentioned above. However, we shall also use the word "idea" as defined above.

It is interesting to note that as far back as 1909 Henry Adams (1949) suggested the possibility of a dynamics of thoughts or ideologies which might have thrown light on the understanding of history. In his suggestion he attempted to construct a dynamics of ideologies along the patterns of dynamics of physical systems. His approach somewhat resembles Haret's approach to social dynamics. We have elsewhere (Rashevsky, 1948, 1951a) pointed out the fundamental weakness of a mere mimicking of physical sciences in dealing with social problems.

In this paper we shall give a very sketchy outline of a dynamics of ideologies based on the mathematical biology of social behavior, as developed previously (Rashevsky, 1951a).

I. *Quantitative Classification of Ideologies.* We shall consider two main qualitative subdivisions of ideologies and within each we shall attempt a quantitative classification.

The first subdivision contains the ideologies concerned with the nature and causes of things. They may be called causational ideologies and spring from attempts to answer the question "why?" All beliefs in deities of various kinds and ideas about their parts in creating and operating the universe belong to this category, and form the first subcategory. To the same category, but to a different subcategory, also belong different scientific ideas concerning the nature of the world. The fundamental difference between the two subcategories is that in the former the acceptance on faith plays a cardinal role, whereas in the latter rational critical analysis is paramount. As we have pointed out elsewhere (Rashevsky, 1951a, chap. xxi) complete freedom from acceptance on faith is a physical impossibility. We cannot verify every statement which we hear and which we intend to use. However, whereas some individuals will accept some statements on faith as a matter of expediency, and will always admit that the statement so accepted may turn out to be wrong, others consider the slightest doubt of certain statements as inadmissible. Herein lies the fundamental difference between the scientific and religious, or rational and arational (*loc. cit.*, p. 170), attitudes. A neurobiological mechanism which explains how the same individual may exhibit both attitudes has been discussed (*loc. cit.*, chap. xxi).

We may call the ideologies of the first subcategory arational, that of the second rational. In this paper we shall restrict ourselves to the former. A theory of the latter will be developed subsequently.

A quantitative classification of that subcategory may be attempted by considering the number of different deities involved in the various beliefs. We thus find a quasi-continuous change going from the innumerable "gods" of the fetishist through the "higher" polytheistic religions to the so-called monotheism of Christianity and Islam. Together with the reduction of the number of deities goes an increase in their abstraction. The process may be readily understood in terms of man's developing knowledge of natural phenomena. In the earliest stages of development man had not learned the fallacy of false generalization. An accidental coincidence between two phenomena led him to consider one phenomenon as the cause of the other. The two phenomena became permanently associated in his mind in a causal relation. A "deification" of different objects or phenomena thus occurs. With progressing knowledge those original associations, which are formed as conditioned reflexes, are gradually dissolved. An individual who learns that he can make a useful tool of which he has full command from a piece of wood which he originally deified will deify it no longer. Thus the total number D of deities is being reduced through knowledge. As man gains more and more control over nature, the concept of deity becomes more and more abstract. Man will not deify an object or phenomenon over which he has control.

Thus one of the first relations which may be looked for in a dynamics of ideas is a relation between D and the total amount K of knowledge. To this end we must first find a proper measure of knowledge, and then establish the law of its variation with respect to time. Possible expressions for K have been tentatively suggested in terms of information theory (Rashevsky, 1950). A development of a theory on that basis may be attempted.

The amount of knowledge is in general distributed non-uniformly in a society. A few individuals have much knowledge, many have little. The general theory of social distributions, developed in chapter viii of *loc. cit.* may be applied to this case. The problem is somewhat simplified by the fact that when two individuals who possess different amounts of knowledge interact, the individual with less knowledge may gain some, while the individual with more knowledge cannot lose any. Thus the function $p(s, s', \Delta)$, where s and s' now denote the amount of knowledge, is such that for $s > s'$, $\Delta = 0$, while for $s < s'$, $\Delta > 0$.

Due to a process of imitation, such as discussed in chapter xiii of *loc. cit.*, it is possible for a social group to adopt, so to speak "officially," an ideology which corresponds to the higher knowledge of the few, although the majority may have a much lesser knowledge. The less enlightened ma-

jority, though adopting the ideology of the more enlightened minority, will likely exhibit some features of behavior which correspond to the lesser level of knowledge. Thus in some Christian countries beliefs in ghosts, mermaids, etc. still persevere among non-educated people.

The opposite effect, however, may also occur: the society as a whole may adopt an idea which corresponds to the lesser knowledge.

As in other cases of imitative behavior, changes from one set of ideologies to another will in general occur discontinuously, even though K may vary continuously.

In a preceding paper (Rashevsky, 1951e, hereafter referred to as *Beliefs*) we suggested a neurobiophysical mechanism for the formation of various beliefs. According to that picture the adoption or non-adoption of some religious belief by a society is determined by the interaction of the process of early conditioning to the belief with the growing tendency to analyze everything critically. The strength of conditioning caused by external factors is determined by the quantity b in equation (2) of *Beliefs*.

Now consider a society composed of n spatially distinct subsocieties which closely interact with each other. Let the belief patterns to which the individuals are conditioned in general be different for the different subsocieties. However, let all those patterns contain one common element. Thus the pattern for the kth subsociety may be represented as consisting of the set of elements

$$B_k = (B_0, B_k^{(1)}, B_k^{(2)} \ldots . B_k^{(r)}), \tag{1}$$

where $B_k^{(i)}$ is in general different from $B_l^{(i)}$, and where B_0 is the common element. Due to the interaction of the subsocieties an individual in the kth subsociety will be principally conditioned to the pattern B_k, but to a lesser extent he will be conditioned to the remaining $(n-1)$ patterns $B_i (i \neq k)$. Let the relative frequency of conditioning to any other pattern be expressed by the fraction $\omega (< 1)$, if the frequency of conditioning to B_k is taken as unity. The greater the interaction of the subsocieties, the greater ω. If $\omega = 1$, this means such a close interaction that each individual is equally exposed to the conditioning to the beliefs of his own subsociety and of any other.

The common element B_0 is conditioned every time any pattern is conditioned. Other elements are conditioned only when their corresponding pattern is being conditioned. Assuming approximate linearity between the final strength of conditioning and the number of repetitions, we find that the strength of conditioning of B_0 will be $1 + \omega(n-1)$ times greater than the strength of conditioning of any other element $B_k^{(i)}$.

In line with the developments presented in *Beliefs*, if we now ask what the probability is that any of the partial beliefs $B_k^{(i)}$ will be abandoned by the society as a whole, we find that this probability depends on the values of the constant b (cf. *Beliefs*) for each of these beliefs. From what we have just said it follows that the value $b^{(0)}$ of b for the belief B_0 is $1 + \omega(n-1)$ times as large as the value $b_k^{(i)}$ of b for any of the other beliefs $B_k^{(i)}$. Thus we have

$$b_0 = [\,1 + \omega\,(n-1)\,]\,b_k^{(i)}\,. \tag{2}$$

In absence of imitation the fraction ν of individuals who do not accept $B_k^{(i)}$ is given by equation (49) of *Beliefs*, into which equation (20) should be introduced for $\bar{\bar{\varphi}}_\infty$. This gives

$$\nu = \frac{1}{2}\left[1 - 2G\left(\frac{\frac{b\epsilon_0}{\gamma}\,e^{-\gamma\tau_0} - \frac{\lambda}{\gamma}\,e^{-\gamma\tau^*}}{a\sqrt{s^2+\sigma_\infty^2}}\right)\right]. \tag{3}$$

If λ/γ is sufficiently large, then for sufficiently small values of b the fraction ν will approach unity. For sufficiently large values of ω and n, the values of $b_k^{(i)}$ may be sufficiently small to insure a non-acceptance of the beliefs; yet the value b_0 may be, because of (2), still so high that belief B_0 is accepted by the society as a whole.

Thus the condition for the survival of the common element of several belief patterns is a sufficiently large number n of subsocieties and a sufficiently strong cultural interaction ω between these subsocieties.

Such conditions are realized by large empires composed of numerous countries with different belief patterns. Let B_k stand for any polytheistic religion. The common element B_0 is the belief in the existence of deity in general. The number of deities and their attributes constitute the elements $B_k^{(i)}$.

The above discussion may throw some light upon the question of why Christianity, which is relatively monotheistic, spread approximately at the time of the largest extension of the Roman Empire. Other ancient empires were also composed of several subsocieties, but the value of n, and possibly of ω, was the largest for the Roman Empire.

We may also ask the question: What are the characteristics necessary to the particular subsociety in which the monotheistic idea first originated? Clearly, that subsociety will have values of $b_k^{(i)}$ which are already rather small compared to $b^{(0)}$. Such characteristics will be present in a subsociety which previously has had cultural contacts with numerous other societies which have different belief patterns. This condition may

be fulfilled by some nomadic societies, or societies that frequently in the past have been temporarily subjugated by others.

Thus it may not have been a mere accident that the earliest concepts of monotheism appeared in the originally nomadic Jewish society, which formed a part of the Roman Empire at the time when the conditions for a wide spread of monotheism became ripe.

We now pass to the second important category of ideologies. It contains ideologies which deal with the question: How shall a society or individual behave? This is essentially an ethical question. There seems to have been a definite correlation between the ethical codes and religious views in the past. This is probably due to the fact that both depend on some common variables—perhaps, for example, the knowledge K.

Quite generally the types of behavior of a society may be described by the behavior matrix (k_{ij}), as introduced in chapter xiv of *loc. cit.* There we have considered, among others, the following two extreme cases: completely egoistic behavior, characterized by the diagonal matrix

$$\begin{matrix} 1 & 0 & 0 & 0 \ldots . \\ 0 & 1 & 0 & 0 \ldots . \\ \ldots\ldots\ldots\ldots\ldots \\ 0 & 0 & 0 \ldots\ldots & 1 , \end{matrix} \tag{4}$$

and completely altruistic behavior, characterized by the matrix

$$\begin{matrix} 1 & 1 & 1 & 1 \ldots . & 1 \\ 1 & 1 & 1 & 1 \ldots . & 1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 1 & 1 & 1 \ldots\ldots\ldots & 1 . \end{matrix} \tag{5}$$

An important type of behavior is represented by the matrix

$$\begin{matrix} 1 & 0 & 0 & 0 \ldots . & 0 \\ 1 & 0 & 0 & 0 \ldots . & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 1 & 0 & 0 & 0 \ldots . & 0 . \end{matrix} \tag{6}$$

In this case every individual is concerned with the satisfaction or well-being of only *one* individual. That individual is arbitrarily numbered "first" in the matrix (6). This type of behavior is found in patriarchal families, where everything is subject to the desires of one man. It is also found in autocratic societies in which the autocrat is almost deified and considers his

ego superior to everybody else's. Such societies should not be confused with autocratic societies in which the autocrat merely exercises authority over the behavior of the rest of the society, but where behavior is directed toward a goal different from the mere increase of personal pleasure for the autocrat.

If, as we have considered previously (*loc. cit.*), the elements k_{ij} of the behavior matrix are functions of the social distance $s_i - s_j$ of the two individuals, then both matrices (4) and (5) may be represented as limiting cases by a properly chosen expression

$$k(s - s'), \tag{7}$$

which gives the k's as functions of $s - s'$. The general form of those functions are such that $k(x)$ has a maximum for $x = 0$, and decreases with increasing absolute value $|x|$ of x, though they are not necessarily symmetric with respect to $x = 0$. They thus have the general form of distribution functions, and it is convenient to introduce the notions of their standard deviations.

If the standard deviation of $k(s - s')$ is very small, the individual is concerned only with himself and with those others who have the same social status s, because $k(0) = 1$. Since the number of individuals is finite, while the range of s is continuous, strictly speaking, no two individuals will have exactly the same s because the probability of this is infinitesimal. Therefore as the standard deviation σ_s of $k(s - s')$ tends to zero the behavior matrix tends to the form (4). On the other hand as $\sigma_s \to \infty$, the behavior matrix tends to the form (5); because for $\sigma_s = \infty$ the individual is equally concerned with all others.

The elements k_{ij} of the behavior matrix may, however, not only depend on the social distance $s - s'$, but also on the geographic distance r. Therefore we may more generally consider the k's as consisting of two factors, k_s and k_r, such that k_s is a function of $s - s'$, while k_r is a function $k_r(r)$ of r. Thus

$$k(s - s', r) = k_s(s - s')\, k_r(r). \tag{8}$$

We shall denote by σ_r the standard deviation of k_r. For very small values of σ_r the individual is concerned only with himself and with his immediate "geographical" neighbors, even if individuals located farther away have a social status close to his own. For very small values of σ_s the individual is concerned only with himself, even if there are other individuals in close geographic proximity.

The derivative with respect to $s - s'$ of the function $k_s(s - s')$ is an even function of $s - s'$, in the sense that it does not change sign when

s and s' are interchanged though, in general,

$$k_s(s - s') \neq k_s(s' - s). \tag{9}$$

The function $k_s(s - s')$ decreases when $|s - s'|$ increases.

The behavior matrix (6) may be considered as a limiting case of a more general matrix

$$\begin{matrix} k_{11} & k_{12} & \ldots & k_{1N} \\ k_{21} & k_{22} & \ldots & k_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ k_{N1} & k_{N2} & \ldots & k_{NN} \end{matrix}, \tag{10}$$

which is characterized by the following relations:

$$k_{i1} > k_{i2} > k_{i3} > \ldots > k_{iN} \tag{11}$$

and

$$k_{1i} = k_{2i} = k_{3i} = \ldots = k_{Ni}. \tag{12}$$

That is, all elements of a column are equal, while in each row they decrease from left to right.

In this general case the maximum concern of each individual is the "first" one, represented by the first column. This first individual also is primarily concerned with himself. But each individual has some concern about himself, which is, however, less than his concern with anyone to the "left," but greater than his concern with anyone to the "right." We have here a case of complete hierarchy. If we consider the social status as decreasing from left to right, then the matrix (10) may be represented for a very large number of individuals by making $k_s(s, s')$ a function $k_s(s')$ only. Thus, for example, we may put

$$k_s(s') = e^{-as'}. \tag{13}$$

In this case the concern of an individual with another depends only on the social status of that other individual. More generally, we may choose $k_s(s - s')$ in such a way that its derivative with respect to $s - s'$ is an odd function of $s - s'$. For example, we may put:

$$k_s = e^{k(s'-s)}. \tag{14}$$

In this case every individual shows more concern for socially higher individuals than for himself and less concern for socially lower individuals. But his concern now depends on his own social status, also. A very small standard deviation σ_s of k_s in this case reduces the behavior matrix to the form (6), while $\sigma_s \to \infty$ again leads to matrix (5). Thus matrix (5) is a limiting case for both cases of even and odd derivatives of $k_s(s - s')$.

However, the transition from an even function to an odd one involves a discontinuity. The two types of behavior must therefore be considered as mutually exclusive different behaviors.

Both types of behavior can, however, be represented in terms of one continuous variable, which we shall call the index of social individualization i. Let i vary in the range from zero to one. For $0 < i < \frac{1}{2}$ let the society behave according to a matrix characterized by the relations (11) and (12), and let σ_s be a continuous monotone increasing function $\sigma_s(i)$ of i, such that $\sigma_s(0) = 0$, $\sigma_s(\frac{1}{2}) = \infty$. For $\frac{1}{2} < i < 1$, let the behavior matrix be characterized by $k_{ii} > k_{il}$ and $k_{ii} > k_{li}$, and let σ_s be a continuous monotone decreasing function of i, such that $\sigma_s(\frac{1}{2}) = \infty$, $\sigma_s(1) = 0$. For $i = 0$ we now have matrix (6), for $i = \frac{1}{2}$ we have matrix (5), and for $i = 1$ we have matrix (4).

The designation "index of social individualization" is justified because for $i = 0$, that is, for matrix (6), the concern of all but one individual is for that one. For $i = 1$ each individual is primarily concerned with himself. For $i = \frac{1}{2}$ each individual is equally concerned with himself and everyone else. For $i \geqslant \frac{1}{2}$ the prevailing idea is of the equality of all individuals. Even though each individual is concerned primarily about himself, there is no hierarchy present as for $i < \frac{1}{2}$. We may, therefore, call the quantity $E = 2i$, defined in the range $0 < i < \frac{1}{2}$, the coefficient of equality.

The desire of an individual to increase, at least to some extent, the satisfaction of another depends also on the similarity or difference of the satisfaction functions of the two individuals. If the satisfaction functions are similar, in other words, if both individuals have similar tastes and behave similarly, the altruistic element is likely to be much more pronounced. For neurobiological reasons discussed in chapter xxii of *loc. cit.* people are apt to dislike behaviors and tastes which are different from their own. To take this situation into account we must introduce a third factor into k, and write, instead of (8),

$$k(s - s', r, l) = k_s(s - s')\, k_r(r)\, k_l(l), \tag{15}$$

where l measures the distance in the functional space between the satisfaction function S_s of the first individual and $S_{s'}$ of the second. The function $k_l(l)$ decreases with l.

If $x_1, x_2 \ldots x_m$ are the m possible variables on which the satisfaction of any individual in a society may depend, then the satisfaction function S of an individual is a function

$$s(x_1, x_2 \ldots . x_m) \tag{16}$$

of those variables. That function may vary from individual to individual, and for some individuals some of the variables may be absent from the argument. Assuming a Euclidean metric in the functional space, we may put:

$$l^2 = \int \dots \int [S_s(x_1, x_2, \dots, x_m) - S_{s'}(x_1, x_2, \dots, x_m)]^2 dx_1 dx_2 \dots dx_m , \tag{17}$$

where the integration is extended over the whole range of the variables.

If the standard deviation σ_l of $k_l(l)$ is very small, then each individual will try to maximize only the satisfaction of such other individuals as have tastes similar to his own. He will be intolerant of any different tastes. If the form of $k_l(l)$ is such that for sufficiently large values of l, $k_l(l) < 0$, then an individual will be so intolerant of others who have sufficiently different tastes that he actually will try to reduce their satisfaction by punishment or in some other way. This is a rather common phenomenon. When $\sigma_l = \infty$ the individual is equally tolerant of any kind of different taste.

Thus choosing even or odd derivatives of $k_s(s - s')$, as well as the values of σ_s, σ_r, and σ_l we can describe quantitatively the types of ethical behavior ranging from an adoration of one's ruler along with complete self denial, through an egoistic assertion of the supremacy of one's own person, to complete altruism. Different degrees of tolerance to different views, neighborliness, and social "snobbishness" are all included in this general representation. Quite generally we may make $k(s - s', r, l)$ a function of the difference between $k_s(s - s', r, l)$ and $k_{s'}(s - s', r, l)$, in a manner similar to the one suggested on page 40 of *loc. cit.*

Strictly speaking, the standard deviations σ_s, σ_r, and σ_l do not determine the situation completely unless the functions k_s, k_r, and k_l are one-parametric. The most direct, but also the most difficult, approach to the problem is to derive the above functions from some neurobiophysical considerations. If the functions thus derived are found to contain more than one parameter each, we still can apply the same method of classification. A situation will now be described not only by the σ's, but by a larger number of parameters. Pending the development of this suggestion we may, as a temporary expedient, *postulate* some simple one-parametric functions. In the following we shall assume that the functions are one-parametric.

Before concluding this section it must be mentioned that other subdivisions of ideas should also be considered in the development of the theory. We may classify ideas of behavior according to the content and

form of the satisfaction functions. Certain variables may be important in one culture and irrelevant in another. This will be reflected in a change of the form of $S(x_1, x_2, \ldots, x_m)$. A variable x_i may be considered as important if $S(x_1, x_2, \ldots, x_m)$ is very sensitive to a change of this variable so that $|\partial S/\partial x_i|$ is very large. A variable x_k is unimportant if $|\partial S/\partial x_k|$ is small. A proper classification could be made first in a purely formal way and later by ascribing definite meaning to the different variables.

Another important criterion for classification may be seen in the part played by the time element. Individuals may strive to maximize their own satisfaction or that of others at the present moment. Or, they may be willing to sacrifice now, either for the benefit of future generations or for the benefit of a personal "hereafter." The second case is represented by some aspects of Christianity. The first may be found in the contemporary socialistic program of the U.S.S.R.

We characterize the different degrees of this behavior by an index ρ, which may, for example, indicate the relative frequency of acts intended for the benefit of a hypothetical "hereafter" and those intended for the gratification of immediate satisfaction. Thus for a mediaeval monk or hermit ρ would be close to unity, while for an epicurean agnostic it would be close to zero.

Last but not least we must consider the classification of scientific ways of thinking, a problem that has already been mentioned. The transition from the Aristotelian to the non-Aristotelian way of thinking is one case at hand. In a previous paper (Rashevsky, 1950) we discussed the "ideal scientist" by considering the maximalization of a certain function of the knowledge K and of possible information H. It may be possible to construct such a one-parametric function of K and H that for different values of the parameter the maximalization leads to a greater or lesser preference for either K or H. In this way a continuous transition from Aristotelian to non-Aristotelian thinking may be quantitatively described.

II. *Effects of Ideologies on Social Structure.* In *loc. cit.* we saw that the distribution function $f(s)$ of the social status s is determined by two other functions. One is the probability that an individual of ability a and social status s will change his social status by an amount Δ when he interacts with another individual of ability a' and social status s'. This function is designated by

$$p(a, a', s, s', \Delta). \tag{18}$$

The other function determines the relative frequency of interaction of two individuals in terms of their respective social statuses. In *loc. cit.* it

was designated by $K(s, s')$. Here we shall denote it by

$$R(s, s'). \tag{19}$$

It is readily seen that both functions depend on, and perhaps are completely determined by, the functions $k_s(s - s')$, $k_r(r)$, and $k_l(l)$. The function $R(s, s')$ is in many cases of the form $R(s - s')$. In that case the smaller its standard deviation σ_R, the more often will interactions occur only between individuals of nearly equal social status. Since the interaction of two individuals usually implies some degree of mutual concern, therefore, in general, σ_R will increase with σ_s and σ_l. We should also expect the standard deviation σ_p of $p(a, a', s, s', \Delta)$ with respect to Δ to decrease with decreasing σ_s, σ_r, and σ_l. A poor or socially low individual interacting with a rich or socially prominent person, who is concerned only with himself, is not likely to improve his social or economic position as a result of such interaction. Also, if r is great and σ_r small, the interaction may be only an indirect one—say, by mail—and is likely to be less effective.

As before, we may again either attempt to derive the functions $p(\Delta)$ and $R(s - s')$ from the functions k_s, k_r, and k_l; or, as a temporary expedient, we may postulate one-parametric functions for $p(\Delta)$ and $R(s - s')$ and postulate simple relations, e.g., linear ones, between σ_p and σ_R on one hand, and σ_s, σ_r, and σ_l on the other. Since the vanishing of either σ_s, σ_r, or σ_l implies vanishing of σ_p and σ_R, therefore the latter two may be linear in each of the former variables, but the three former should enter in the form of a product. Thus we may put

$$\sigma_R = C_R \sigma_s \sigma_r \sigma_l , \tag{20}$$

$$\sigma_p = C_p \sigma_s \sigma_r \sigma_l , \tag{21}$$

where C_R and C_p are constants.

To include the possibility of pure chance we may put, generally,

$$\sigma_R = C_R \sigma_s \sigma_r \sigma_l + \sigma_R^* (N_0, \eta_i) ; \tag{22}$$

$$\sigma_p = C_p \sigma_s \sigma_r \sigma_l + \sigma_p^* (N_0, \eta_i) ; \tag{23}$$

where σ_R^* and σ_p^* depend on the total population N_0 and on other parameters η_i, such as area occupied by the population, means of communications, etc.

If we can solve the functional equation (21) of chapter viii of *loc. cit.* then we can express $f(s)$ in terms of the parameters σ_s, σ_r, and σ_l which determine the classification of ideas.

Again it may be possible to roughly classify all functions $f(s)$ by their

standard deviations. The problem therefore arises as to whether it is possible to express approximately the standard deviation of $f(s)$ in terms of the standard deviations σ_R and σ_p without actually determining $f(s)$ completely.

III. *Fundamental Equation of the Dynamics of Ideologies.* Let us now apply the reasoning of chapter xxii of our book (Rashevsky, 1951a) to the case of n mutually exclusive behaviors. Following a suggestion of H. D. Landahl's (1941) which we have developed somewhat further in a recent paper (Rashevsky, 1951c), we now shall consider the n quantities

$$\varphi_i = \epsilon_i - \frac{\Sigma \epsilon_k}{n-1} \quad (k \neq i) . \tag{24}$$

The quantity φ_i now determines the probability that behavior R_i will be exhibited. If $\varphi_i = 0$, then R_i is as probable as any one of the other $n-1$ behaviors. It is then equal to $1/n$. As $\varphi_i \to \infty$, the probability P_i of behavior R_i tends to 1. If $\varphi_i < 0$, and tends to $-\infty$, then $P_i \to 0$. The actual expressions for P_i have been given elsewhere (Rashevsky, 1951c).

If n is very large we may substitute integration for summation. Let η be a continuously varying parameter, each value of which corresponds to a particular choice of behavior. For example, in some cases η may be identified with the index of social individualization i. Or η may stand for σ_r or σ_l. Let η vary in the interval (η_1, η_2). For every value of η there is a corresponding value $\epsilon(\eta)$. Denoting by η_m the value of η for which $\varphi(\eta)$ has a highest value, we now put

$$\varphi(\eta_m) = \epsilon(\eta_m) - \bar{\epsilon}(\eta) \tag{25}$$

where

$$\bar{\epsilon}(\eta) = \frac{1}{\eta_2 - \eta_1} \int_{\eta_1}^{\eta_2} \epsilon(\eta)\, d\eta \tag{26}$$

and the integration is taken over all values of η except η_m.

If the total ideology or behavior is characterized by several parameters η_i, then the ϵ's and φ's are functions of all of these parameters. Without loss of generality we shall, for simplicity, confine ourselves here to the one-dimensional case.

With the definitions given above we now may apply the same method as for the choice between two different ideas or behaviors. The probability $P(\eta)$ of an ideology η may now be given by

$$P(\eta) = 1 - e^{-\kappa\varphi(\eta)} \tag{27}$$

where σ is a constant, or by:

$$P(\eta) = G\left[\frac{\kappa_n - \varphi(\eta)}{\kappa}\right], \tag{28}$$

where κ_n is as defined in Rashevsky (1951c).

It may at first glance seem more appropriate to consider different values of η as corresponding to different intensities of the same behavior. However, considering the actual inability of most persons to see essential similarities between only slightly different situations and to realize that the differences are of quantitative rather than qualitative nature, it seems justifiable to use the approach suggested above. A good example of this is the bitter antagonism shown each other by rival political parties whose political ideas actually differ but slightly.

The approach suggested above is justified moreover from a neurobiophysical point of view. The mathematical biophysics of discrimination of intensities (Rashevsky, 1948, chap. xxiii) leads to the concept that qualitatively similar stimuli of different intensities result in the excitation of spatially different regions in the central nervous system. Thus even several different intensities of the same type of behavior lead neurobiophysically to the same model as do several spatially different mutually exclusive stimuli.

The neurobiophysical picture will, moreover, help us to establish the equations which govern the shifts from one value of η to another, equations which were very tentatively suggested in our book (Rashevsky, 1951a).

Consider a structure consisting of n parallel, mutually cross-inhibiting pathways (Rashevsky, 1948, chap. xxxi). Let the ith pathway be excited at the afferent end by an amount ϵ_i of the excitatory factor. Let the parameters of the system be such that when all ϵ_i's are equal, all the responses at the efferent ends are just inhibited. If a single ϵ_k exceeds the average value $\bar{\epsilon}$ of the other ϵ_i's, then the reaction R_k is produced. For very large n this picture leads to the equations (25)—(28), if each reaction R_i represents the behavior corresponding to a particular value η_i in our quasi-continuous scale.

The excitations ϵ_i's are considered to be of a central origin, due to some endogenous factors, just like ϵ_1 and ϵ_2 in the case of only two alternative behaviors. However, let the centers which produce the ϵ's be connected by cross-excitatory pathways along which the intensity of excitation decreases with distance, as was studied previously (Rashevsky, 1948, pp. 381–82; Landahl, 1938). Thus if a given ϵ_i is increased, then the ϵ's in

the immediately adjacent centers will also increase, this increase diminishing with distance from the given center. For definiteness let us consider the case in which the decrease of the cross-excitation with distance is exponential. If we visualize all the n pathways arranged uniformly in a linear array, then the actual distance measured along this array between two points will be proportional to the difference $\eta_i - \eta_k$,where η_i and η_k are the values of η which correspond to those two points respectively.

From the above it follows that if all ϵ's remain equal to each other, except one, $\epsilon(\eta_m)$, which for any reason increases, then the values of $\epsilon(\eta)$ in the neighborhood will also increase, so that $\epsilon(\eta)$ will now be represented by

$$\epsilon(\eta) = \epsilon_0 + \epsilon(\eta_m)\, e^{-\mu^2|\eta-\eta_m|} \tag{29}$$

where ϵ_0 is the level of excitation of the other centers and μ^2 is a constant. If μ^2 is sufficiently large, the distribution of $\epsilon(\eta)$ will represent a sharp peak in the neighborhood of η_m.

From equations (25), (26), and (27), together with equation (29), it follows that the probability $P(\eta)$ is highest for $\eta = \eta_m$. Without imitation the largest percentage of all individuals will exhibit it. In the presence of imitation an absolute majority will exhibit behavior η_m. The actual percentage of that majority is obtained from the previously established equations (Rashevsky, 1951a), in which we substitute $\varphi(\eta_m)$ for ϕ.

As we have seen, the choice of a particular behavior η_i results in a particular social distribution function $f_i(s)$. The latter in turn determines the rate at which the chosen behavior will be either enforced or weakened by learning. Whether it is reinforced or weakened depends (Rashevsky, 1951a, chap. xxii) on the sign of the expression

$$\int_0^1 f_i(s)\, \lambda(s)\, ds\,, \tag{30}$$

where $\lambda(s)$ is positive for those values of s for which the behavior results in satisfaction, and negative for those values for which the behavior is unsatisfactory. We may, for example, put:

$$\lambda(s) = \alpha^2 s - \beta^2\,, \tag{31}$$

where α and β are constants. If $\lambda(s)$ is to be positive for $s = 1$, then $\alpha^2 > \beta^2$. If $\alpha^2 = 2\beta^2$, then

$$\lambda(s) = \alpha^2 (s - \tfrac{1}{2}) \tag{32}$$

so that $\lambda(s) = 0$ for $s = \frac{1}{2}$. It seems more plausible that $\lambda(s) = 0$ for $s \ll \frac{1}{2}$, which implies $\alpha^2 \gg \beta^2$. In line with the discussions of the

satisfaction function (Rashevsky, 1951a) it may be more logical to put

$$\lambda(s) = \log a^2 s \tag{33}$$

with

$$a^2 \gg \tfrac{1}{2} . \tag{34}$$

Generalizing the concepts developed in chapter xxii of our book (Rashevsky, 1951a) we may consider that a negative total satisfaction reduces the value of $\varphi(\eta_m)$ by reducing $\epsilon(\eta_m)$. Due to the cross-excitatory connections, this will also reduce the adjacent $\epsilon(\eta)$'s.

Without loss of generality we may put $\eta_m = 0$ and let η vary in the interval (0, 1). Let the society initially have a preference for behavior and ideology $\eta = 0$, and let the value of $\epsilon(\eta_m)$ at the time $t = 0$ of adoption of that ideology be

$$\epsilon(\eta_m) = \xi . \tag{35}$$

If that ideology or behavior results in an average dissatisfaction, then ξ will decrease with time, and at any time $t > 0$ we shall have

$$\epsilon(\eta) = \epsilon_0 + \xi(t)\, e^{-\mu^2 \eta} . \tag{36}$$

Eventually $\xi(t)$ will become negative and $\epsilon(\eta)$ will have its smallest value at $\eta = 0$ and largest at $\eta = 1$. That largest value is equal to

$$\epsilon_0 + \xi(t)\, e^{-\mu^2} < \epsilon_0 . \tag{37}$$

The value of $\varphi(1)$ now increases. At any moment t it is given, according to equations (25), (26), and (37) by

$$\varphi(1) = \xi(t)\, e^{-\mu^2} - \xi(t) \int_0^1 e^{-\mu^2 \eta} d\eta = -\frac{\xi(t)}{\mu^2} + \frac{\mu^2 \xi(t) + \xi(t)}{\mu^2}\, e^{-\mu^2} . \tag{38}$$

If, as we shall assume, $\mu^2 \gg 1$, then, approximately,

$$\varphi(1) = -\frac{\xi(t)}{\mu^2} > 0 . \tag{39}$$

As soon as $\varphi(1) > \varphi^*$, where φ^* is the threshold determined by the slope of the $F(\psi)$ curve and the slope of the $a\psi$ line (Rashevsky, 1951a, chap. xii), society will switch into ideology and behavior $\eta = 1$. This will happen when

$$-\xi(t) = \mu^2 \varphi^* . \tag{40}$$

If this behavior results in a net dissatisfaction, then the value of $\epsilon(1)$, which at the moment of the switch was $\epsilon_0 + \xi e^{-\mu^2} \sim \epsilon_0$, will begin to decrease, and at every moment t be given by

$$\epsilon_0 + \zeta(t), \qquad \zeta(t) < 0, \tag{41}$$

if we now put as the new origin of time t the moment when ideology $\eta = 1$ is adopted.

Because of the excitatory cross connections this will result in an addition to every $\epsilon(\eta)$ of the (negative) amount

$$\zeta(t)\, e^{-\mu^2(1-\eta)}. \tag{42}$$

At the same time the original inhibition, which reduces ϵ, is now decreasing, due to "forgetting" (Rashevsky, 1951a, p. 180). This decrease may be assumed to proceed exponentially with time. Therefore, we now have

$$\epsilon(\eta) = \epsilon_0 - \mu^2\varphi^* e^{-\gamma t} e^{-\mu^2\eta} - \zeta(t)\, e^{-\mu^2(1-\eta)}. \tag{43}$$

The function $\epsilon(\eta)$ has a maximum at

$$\eta_{\max} = \frac{1}{2\mu^2}[\mu^2 - \gamma t - \log \zeta(t) + \log \mu^2\varphi^*. \tag{44}$$

This maximum is somewhere between $\eta = 0$ and $\eta = 1$, and shifts with time. It is readily seen that it also becomes more pronounced. By introducing expression (44) into (43), and subtracting from this the integral of (43) with respect to η from 0 to 1, we find the expression for $\varphi(\eta_{\max})$ as a function of t. Setting this equal to φ^* we find the moment t_1 at which the society will switch to a new behavior. By introducing this value of t_1 for t into expression (44) we find the $\eta_{\max}$ which describes the new ideology and behavior. From there on we again repeat the procedure, remembering that now expression (42) is also to be multiplied by $e^{-\gamma t}$. In addition, since the new behavior lies between $\eta = 0$ and $\eta = 1$, we shall have to study both branches of the curve (29).

We thus have given a possible analytical treatment to the problem outlined on pp. 180–81 of our book (Rashevsky, 1951a).

There is, however, another possible way for society to shift from one ideology or behavior to another. If expression (30) is positive, which is equivalent to a positive average satisfaction, then the value of $\varphi(\eta)$ for the chosen behavior η will increase. In the absence of other factors this would make the behavior η permanent. The fraction of individuals accepting it would tend, with time, to one.

We must remember, however, that the quantities i, σ_r, and σ_l, which may stand for η, vary in general with increasing intellectual development of society. They will depend on the knowledge K, and especially on the amount of communication transmitted from individual to individual in a society. The greater the density of population, the more likely is an individual to come in contact with others and, therefore, the more likely he

is to know other individuals well. This better knowledge is likely to increase σ_s, σ_r, and σ_l. Thus, even in the absence of any satisfaction or dissatisfaction, that is, even if expression (30) is zero, a slow continuous shift of η will occur. Mathematically this can be taken into account by making the quantities which correspond to η_m in equation (29) be slowly varying functions of time at every step of the previous discussion.

Thus when a particular ideology or behavior η is chosen, two things may happen:

If the chosen behavior results in discontent, then it will eventually be abandoned for one which is remote on the scale of behavior. The transition to the new behavior or ideology will be sudden when the threshold φ^* is exceeded.

If the chosen behavior is, on the average, emotionally neutral, or even pleasant, then a slow shift in the chosen η will occur. From a behavior characterized by a value of η_m a change will be made to one characterized by $\eta_m + \Delta\eta$, where $\Delta\eta$ is a small increment. The change will be, strictly speaking, discontinuous, in the sense that the majority of individuals who exhibited behavior η will suddenly begin to exhibit behavior $\eta_m + \Delta\eta$. Because of the smallness of $\Delta\eta$, however, the total process will appear to be continuous.

From equations (30)—(34) we see that the total dissatisfaction increases with increasing steepness of the distribution function $f(s)$. If, as a measure of that steepness, we take the standard deviation σ_f of $f(s)$, then we see that for small values of σ_f sudden changes from one behavior to another will occur. For sufficiently large values of σ_f the transitions will be continuous. This agrees with the results at which we arrived previously (Rashevsky, 1951a, chap. xxvi) by a different method.

The previously developed equations which govern imitative behavior, together with equations (25) and (26), give us the percentage X of the individuals who exhibit the chosen behavior (Rashevsky, 1951d). The percentage Y now includes all the individuals who exhibit all the non-η_m behaviors. It is interesting to inquire how the other behaviors will be distributed within the fraction Y. It is natural to assume (and an attempt to prove it may be worthwhile) that the relative portion of individuals among the fraction Y who exhibit different non-η_m behaviors is proportional to the corresponding values of $\epsilon(\eta)$ for $\eta \neq \eta_m$. Or it may be that of all the behaviors $\eta \neq \eta_m$ the one is preferred which has the highest value of $\epsilon(\eta)$ and the preference is determined by a mechanism similar to the preference for η_m among all η's.

Before any applications of the theory outlined can be made it will have

to be developed considerably. By way of an illustration as to what eventually may be expected from such a theory we may very tentatively suggest the following interpretation of some historical processes.

Toward the end of the Roman Empire behavior was characterized by a matrix rather similar to (6). The index i of individualization was near zero for the majority of the population. The subjugated people and the vast number of slaves worked for their masters, whose number was relatively small.

Due to factors discussed in section I of this paper the ideology of monotheism then spread in the form of Christianity. If we classify the latter in our system of ideologies, we see that with its idea of the brotherhood of man it corresponds essentially to matrix (5), with $i = \frac{1}{2}$ and $\sigma_r = \sigma_l = \infty$. Another characteristic of Christianity was the development of asceticism with belief in reward in the hereafter; hence with a value of ρ near unity. Of the above characteristics only the second prevailed in practice. Soon after the beginning of Christianity the idea of universal brotherhood became non-existent in practice. The large value of ρ is likely to have been caused by the highly unsatisfactory position of large oppressed masses. Those masses could relieve their dissatisfaction only in one of two ways: by gaining actual control of the society and improving their situation, or by reducing their dissatisfaction through reducing their desires and through hopes for a better hereafter. Physically the two ways lead to very different results, but psychologically the results are similar [cf. P. Sorokin's definition of freedom (1937)].

Thus at the beginning of the Middle Ages ideology and behavior were characterized by $i = 0$, $\sigma_r \ll 1$, $\sigma_l \ll 1$, $\rho \sim 1$. This, as we have seen, is likely to lead to a very steep $f(s)$, that is, to a very small σ_f. Though no quantitative data on $f(s)$ in the Middle Ages are available, the social stratification is known to have been very pronounced. This bears out our conclusion.

But a small value of σ_f should have resulted in an eventual sudden transition to an ideology with $i = 1$, $\rho \sim 0$, $\sigma_r \sim 1$, $\sigma_l \sim 1$. This sudden transition should, however, have been preceded by a relatively long period with an apparent lack of change. As we have seen before (Rashevsky, 1951a) the gradual shift in φ, or, in our case, in $\varphi(\eta)$ results in a parallel displacement of the curve $F(\psi)$ (Fig. 1) to the right. If the slope of the straight line is very small, or if the initial $\bar{\varphi}$ is sufficiently positive so that the intersection with the $F(\psi)$ curve is very near the asymptote, then a displacement of that curve will not materially alter the value of X, until the curve reaches a position somewhere in the neighborhood of the broken

line. From there on X begins to decrease appreciably, until the position indicated by the dotted line is reached. At that point X changes discontinuously to a value which is near zero.

The period between the position of the full line and that of the broken line may be considered to correspond to the Middle Ages proper. The position of the broken line may be considered as the beginning of the Renaissance, while the position of the dotted line may be considered to correspond to the quasi-discontinuous revolutionary changes of the end of the 18th Century and the beginning of the 19th. Those changes, as we have seen, should have resulted in ideologies which correspond to

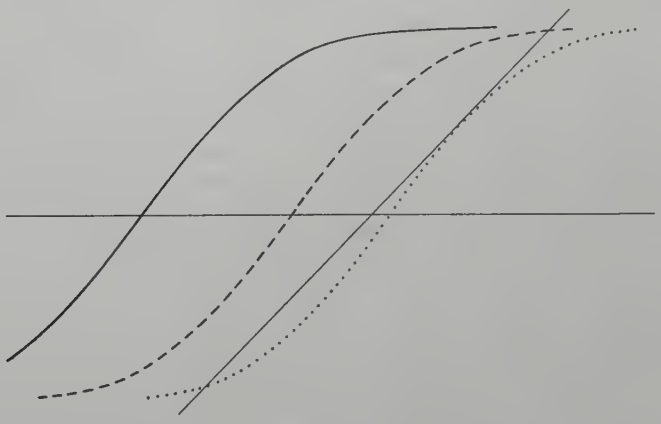

FIGURE 1

$\rho \ll 1$, $i \sim 1$. But $i = 1$ corresponds to matrix (4), which is characteristic of extreme individualism. Economically matrix (4) corresponds to what is usually designated as "free enterprise" (Rashevsky, 1951a, pp. 146–7). A small ρ means, in common parlance, a more materialistic outlook on life, with the idea of any "hereafter" taking a very subordinate place. But such changes have been characteristic of the late 18th and early 19th centuries.

While the values of i, σ_r, and σ_s affect the form of $f(s)$, they do not determine $f(s)$ completely. A number of other factors also affect $f(s)$, especially if s denotes the economic status. Therefore after the idea of individualism established itself in different countries, its further fate varied from country to country. In western Europe with a high density of population and high industrialization σ_f was relatively large. Hence further shifts in i and in the other parameters took place quasi-continuously. The results of previous studies (Rashevsky, 1951a, chapters xiv, xv)

indicate that a minimum total dissatisfaction is likely to occur in a society characterized by matrix (5), that is, by $i = \frac{1}{2}$. Hence we should expect a quasi-continuous shift toward such a behavior in western European countries. On the contrary, in a country with little industrialization and a high degree of total dissatisfaction, like czarist Russia, a sudden transition to extreme socialism was to be expected.

Some of the conclusions outlined here may in principle be checked by gathering qualitative or semi-qualitative data on socio-historical developments of the type, for example, compiled by P. Sorokin in his *Social and Cultural Dynamics* (1937). Among other things Sorokin attempts to follow by centuries the fluctuation of such qualities as the relative frequency of appearance of religious and secular subjects in art, or the relative frequency of description of different social classes in art. Such figures, inexact as they are, may be indications of the relative values of X and Y, where X may stand, for example, for the fraction of individuals who exhibit characteristically mediaeval behavior, with $\sigma_r \ll 1$, $\sigma_l \ll 1$, $\rho \sim 1$; while Y represents the fraction of individuals who exhibit the opposite behavior which is characteristic of modern times. Before making use of such data, even if they were accurate, an important question must be decided. Do the values of X and Y as given by such data represent the *actual* values for the society at a given period, as determined by mass behavior and given by $\varphi + \psi$, or do they reflect the innate values of φ, which would manifest themselves in the absence of imitation? At first glance the former possibility seems to be the natural one. However it is possible that the artist reflects the inner tendencies of his fellow men more than their overt behavior. As we have seen (Rashevsky, 1951a), the average $\bar{\varphi}$ may have already shifted well toward the opposite behavior, yet the majority of society may still exhibit the given behavior.

In order merely to illustrate what *could* be done with such data, let us assume the second situation. In that case X is determined not by the intersection of the curve $F(\psi)$ with the straight line $a\psi$, but by the ordinate of the curve $F(\bar{\varphi})$, where $\bar{\varphi}$ is the average value of φ at a given moment. We have

$$X = \tfrac{1}{2}\,[\,1 + F\,(\bar{\varphi})\,] \tag{45}$$

and, approximately,

$$F\,(\bar{\varphi}) = G\left(\frac{\bar{\varphi}}{\kappa}\right), \tag{46}$$

because the quantities κ_m (Rashevsky, 1951c) are very small for large n and may be neglected.

Let us assume that $\bar{\varphi}$ varies linearly with time:

$$\bar{\varphi} = v\,(t - t_0)\,, \tag{47}$$

and that positive φ corresponds to "mediaeval" behavior. We now introduce the new variable

$$\bar{\varphi}' = \frac{\bar{\varphi}}{v} = (t - t_0)\,. \tag{48}$$

Then, putting

$$\kappa' = \frac{\kappa}{v} \tag{49}$$

we have, from (46) and (47):

$$F\,(t - t_0) = G\left(\frac{t - t_0}{\kappa'}\right). \tag{50}$$

From the values of X at various times, we find

$$F\,(t - t_0) = 2X - 1 \tag{51}$$

and then we may calculate κ' and t_0. In Figure 2 we have plotted the values of F computed from Sorokin's data on the relative percentage of religious and secular art for central Europe (Sorokin, 1937, Vol. I, p. 381), marked by squares; data on same frequencies for Europe as a whole (*ibid.*, p. 382), marked by asterisks; and data on relative frequencies of representation of different social classes (*ibid.*, p. 486), marked by triangles. The full line represents the G-curve, which fits those data as well as is possible in view of their inaccuracy. Sorokin's data (Fig. 2) gives us, if we use 100 years as the unit of time, $\kappa' = 0.5$, $t_0 \sim 1650$ A.D. We see that $\bar{\varphi}' = 0$ and hence $\bar{\varphi} = 0$ at about 1650 A.D. Since the actual revolutionary transition toward negative φ's occurred at the end of the Eighteenth Century (French Revolution), therefore this supports our assumption that Sorokin's data gives us $\bar{\varphi}$ rather than $\bar{\varphi} + \psi^*$. For $t > t_0$ the F curve, as seen from (50), is merely displaced to the right. At approximately 1800 A.D. it must have been in the position shown by the broken line of Figure 2. At that time the sudden transition occurred; hence it was then that the straight line $a\psi$ became tangent to the curve. Since the transformation (48) does not alter the ordinate of the point of tangency, but merely changes the units of φ so that numerically φ becomes equal to $t - t_0$, by drawing a tangent to the F curve in Figure 2 from the point $t - t_0 = 0$, we find the value of F and hence of X at the moment of the revolution. The value is approximately $X = 90\%$, which is quite plausible. After the revolution we have $X \sim 0.02$, $Y \sim 0.98$, which is again plausible. In principle those values could be determined by appropriate quantitative historical research.

One more illustration may be given. During the Middle Ages X was of the order of 98%, Y of the order of 2%. As we have seen elsewhere (Rashevsky, 1951b), the scientific and technological productivity of a society should be approximately proportional to Y. If this is the case we should expect approximately a 50-fold increase in that productivity. Since the French Revolution P. Sorokin (1937) finds the approximate number of discoveries and inventions made in western Europe between 600 A.D. and 1800 A.D. to be 3,000. From 1800 A.D. to 1900 A.D. he finds that number to be approximately 8,500. Per unit time this is an increase of 34 times.

We again must emphasize that all the above is offered only as an illus-

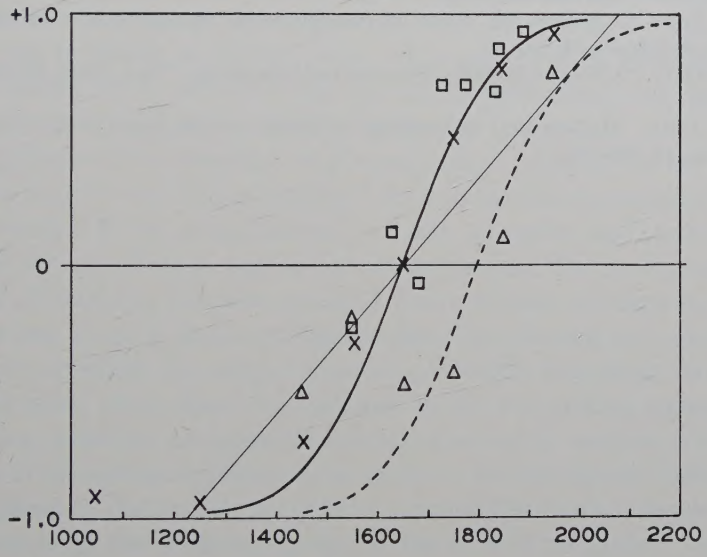

FIGURE 2

tration of how mathematical reasoning could be applied to historical processes. Before this actually can be done a great deal of further development of the theory will have to be made. Moreover only if such a development stimulates quantitative historical research, which now is practically nonexistent, can any applications be made.

Important quantitative questions arise in connection with comparative studies of different cultures. Why did the West develop earlier industrially than the Orient? What factors determine the absolute rates of development? Those questions we do not attempt to discuss here, but hope to do so in future publications.

The author is indebted to Dr. George Karreman for a critical discussion of the paper.

LITERATURE

Adams, Henry. 1949. *The Degradation of the Democratic Dogma.* New York: Peter Smith.

Jacobson, Edmund. 1930. "Electrical Measurements of Neuromuscular States During Mental Activity." *Am. Jl. Physiol.*, **95**, 694–702; 703–12.

———. 1931. "Electrical Measurements of Neuromuscular States During Mental Activity." *Am. Jl. Physiol.*, **96**, 115–21; **97**, 200–9.

Rashevsky, N. 1948. *Mathematical Theory of Human Relations, An Approach to a Mathematical Biology of Social Phenomena.* Bloomington (Indiana): The Principia Press.

———. 1950. "Some Bio-Sociological Aspects of the Mathematical Theory of Communication." *Bull. Math. Biophysics*, **12**, 359–78.

———. 1951a. *Mathematical Biology of Social Behavior.* Chicago: University of Chicago Press.

———. 1951b. "Suggestions for a Mathematical Biology of Some Cultural Developments." *Bull. Math. Biophysics*, **13**, 51–60.

———. 1951c. "A Note on the Theory of Communication through Social Channels." *Bull. Math. Biophysics*, **13**, 139–46.

———. 1951d. "A Note on Imitative Behavior and Information." *Bull. Math. Biophysics*, **13**, 147–51.

———. 1951e. "Mathematical Biosociology of Beliefs and Prejudices." *Bull. Math. Biophysics*, **13**, 289–301.

AWARDS FOR POST-DOCTORAL STUDY IN STATISTICS AT THE UNIVERSITY OF CHICAGO

The Committee on Statistics (a department) of The University of Chicago has established, under a five-year grant from the Rockefeller Foundation, a program of Post-doctoral Awards to provide training and experience in statistics for scholars whose main interests lie outside that field. There will be three Awards per year, to holders of the doctorate or equivalent in the biological, the physical, and the social sciences. Each Award will be $4000 or slightly more, office space will be provided, and $600 to $1000 will be available for clerical, computational, and research assistance. There will be no tuition charges.

The purpose of the Awards is to give statistical training to a few scientists who may be expected to employ it both to the direct advance of their specialties and to the enlightenment of their colleagues and students by example, by consultation, and by formal instruction. The development of the field of statistics has been so rapid that problems of communication are a serious obstacle to its full exploitation. The amount and quality of instruction available to young students is constantly increasing, but there is a real need, which these Awards seek to fill, for making appropriate instruction available to already established scientific workers who give promise of immediate applications of statistics to their special fields.

Recipients of the Awards must have received the doctor's degree prior to commencing the program, except in the case of recognized research workers whose experience and accomplishments are clearly the equivalent. Candidates whose mathematical preparation includes less than the usual sophomore year of calculus, or its equivalent, will not ordinarily be considered, but previous training in statistics is not required or expected. Candidates having under way research programs in their own fields will be preferred, and the department of The University of Chicago concerned with a candidate's specialty will be asked to participate in evaluating his application. Recipients must spend eleven months studying statistics at The University of Chicago, and will be expected to pursue a number of regular courses.

Applications, or requests for further information, should be sent to: Committee on Statistics, The University of Chicago, Chicago 37. Applications for the academic year 1952–53 should arrive by April 1, 1952.